



**VERİ AMBARI VE OLAP TEKNOLOJİLERİNDEN  
YARARLANILARAK KARAR DESTEK AMAÇLI  
RAPORLAMA ARACI GERÇEKLEŞTİRİMİ**

**IMPLEMENTATION OF A REPORT TOOL FOR DECISION  
SUPPORT SYSTEM WITH DATAWAREHOUSE AND OLAP  
TECHNOLOGIES**

**SELDA DÜZGÜNOĞLU**

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin  
Bilgisayar Mühendisliği Anabilim Dalı İçin Öngördüğü

**YÜKSEK LİSANS TEZİ**

olarak hazırlanmıştır.

2006

Fen Bilimleri Enstitüsü Müdürlüğü'ne,

Bu çalışma jürimiz tarafından **BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI'nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Üye (Danışman) :.....  
Prof. Dr. Adnan YAZICI

Başkan (Danışman) :.....  
Prof. Dr. Ünal YARIMAĞAN

Üye :.....  
Dr. İbrahim ZOR

Üye :.....  
Dr. Harun ARTUNER

Üye :.....  
Dr. Mustafa EGE

ONAY

Bu tez .... / ..... / ..... tarihinde Enstitü Yönetim Kurulunca belirlenen yukarıdaki jüri üyeleri tarafından kabul edilmiştir.

..... / ..... / .....

Prof.Dr. Ahmet R. ÖZDURAL  
FEN BİLİMLERİ ENSTİTÜSÜ MÜDÜRÜ

# VERİ AMBARI VE OLAP TEKNOLOJİLERİNDEN YARARLANILARAK KARAR DESTEK AMAÇLI RAPORLAMA ARACI GERÇEKLEŞTİRİMİ

**Selda Düzgünoğlu**

## ÖZ

Günümüzde veri ambarı ve OLAP(ÇİAİ : Çevrimiçi Analitik İşleme) teknolojilerinin kullanımı gün geçtikçe yaygınlaşmaktadır. Günlük işlerin takip edilmesi amacıyla hazırlanmış otomasyon yazılımları az seviyede raporlama içeren ve yüksek seviyeli hareket işleme amaçlı tasarlanmış işletimsel sistemlerdir (ÇİHİ : Çevrimiçi Hareket İşleme). Veri ambarı uygulamaları, analiz ve raporlama amaçlı tasarlanmış ve analitik işlemleri destekleyen sistemlerdir. İşletimsel sistemlerle üretilmiş verilerin veri ambarı ortamında varsa mevcut hatalarından ayıklanarak ve gerekli veri dönüştürmeleri yapılarak özet bilgilerin saklanması sağlanır. Toparlanmış özet veriler üzerinde ÇİAİ teknolojisi kapsamında tanımlanan küp veri yapıları ile karar destek bağlamında sonuç üretebilecek sorgular hazırlanır. Bu yaklaşımların kullanılması gelecek için alınacak kararların belirlenmesinde büyük önem taşımaktadır.

Tez kapsamında, karar destek amaçlı raporların üretilmesini sağlayacak kolay kullanımlı bir raporlama aracı tasarlanmıştır. Sistemde bulunan işletimsel veriler veri ambarı yapısında tekrar organize edilmiş ve veri küpleri tanımlanarak istatistiksel raporların hazırlanmasına çalışılmıştır. Karar destek kapsamında raporlama aracı geliştirilerek, mevcut teknolojilere hakim olmayan son kullanıcıların rahatlıkla kullanabileceği, hızlı sorgulama yapabilen, karar alımına yardımcı bir uygulama olanaklı kılınmıştır.

**Anahtar Sözcükler:** Veri ambarı, OLAP, ÇİAİ, Karar Destek Sistemi

Danışman: Prof.Dr. Adnan YAZICI, ODTÜ, Bilgisayar Mühendisliği Bölümü

Danışman: Prof.Dr. Ünal YARIMAĞAN, Hacettepe Üniversitesi, Bilgisayar Mühendisliği Bölümü

# **IMPLEMENTATION OF A REPORT TOOL FOR DECISION SUPPORT SYSTEM WITH DATAWAREHOUSE AND OLAP TECHNOLOGIES.**

**Selda Düzgünoğlu**

## **ABSTRACT**

Today, using of data warehouse and OLAP technologies become widespread. Operational softwares which are designed for daily operations include low level reporting and high level transactions. Those systems are transactional and called as OLTP systems. Data warehouse applications are designed for analyze and reporting and support analytical operations. Data which is produced from the operational systems is extracting, transforming and loading into the data warehouse environment. The aggregated values of the data are mainly stored in the data warehouse. By defining multidimensional views named 'cube' from the aggregated data by using OLAP technologies, queries are processed for decision support. Using these approaches carries great importance to find decisions for the future.

In my thesis, a user friendly report generation tool for decision support is designed. The transactional data into the database is reorganized and after transformations it is loaded into the data warehouse environment. Statistical reports are generated by using multidimensional data cubes. Including decision support, a reporting tool is developed. So that, users who don't know existing technologies can use the tool easily. And the reporting tool also reduces query result time by using OLAP technologies. An application which supports decision support is developed in my thesis.

**Keywords:** Data warehouse, OLAP, Decision Support System

Advisor: Prof.Dr. Adnan YAZICI, ODTÜ, Department of Computer Engineering

Advisor: Prof.Dr. Ünal YARIMAĞAN, Hacettepe University, Department of Computer Engineering

## **TEŞEKKÜR**

Yazar, bu çalışmanın gerçekleşmesinde katkılarından dolayı, aşağıda adı geçen kişi ve kuruluşlara içtenlikle teşekkür eder.

Sayın Prof.Dr. Adnan Yazıcı (Tez Danışmanı), yazara kendisi ile çalışma fırsatı vermiş, tez konusunun belirlenmesinde, tezin gerçekleştiriminde ve tez metninin oluşmasında değerli fikir ve tecrübeleri ile yön göstererek, tez süresinde danışmanlık yapmışlardır.

Sayın Prof.Dr. Ünal Yarımağan (Tez Danışmanı), yazara kendisi ile çalışma fırsatı vermiş, tez konusunun belirlenmesini sağlamış, tezin gerçekleştirimi sırasında yol gösterici olmuşlardır.

Sayın Dr. Mustafa Ege, tez metnini inceleyerek biçim ve içerik bakımından son şeklini almasında katkıda bulunmuşlardır.

Sayın Dr. İbrahim Zor, tez metnini inceleyerek biçim ve içerik bakımından son şeklini almasında katkıda bulunmuşlardır.

Sayın Dr. Harun Artuner, tez metnini inceleyerek biçim ve içerik bakımından son şeklini almasında katkıda bulunmuşlardır.

Sayın Ebru Aylin Bayrak, tezin gerçekleştirimi sırasında destek olmuş ve tez metnini inceleyerek biçim ve içerik bakımından son şeklini almasında katkıda bulunmuşlardır.

Sayın Cemal Koplay, tez metnini inceleyerek biçim ve içerik bakımından son şeklini almasında katkıda bulunmuşlardır.

Tepe Teknolojik Sevisler A.Ş., Hastane Bilgi Yönetim Sistemi bilgi birikimini yazarla paylaşarak veri ambarı tasarımında yönlendirici olmuştur.

## İÇİNDEKİLER DİZİNİ

|  |      |
|--|------|
| ÖZ .....   | i    |
| ABSTRACT .....   | ii   |
| TEŞEKKÜR.....  | iii  |
| ŞEKİLLER DİZİNİ .....  | viii |
| ÇİZELGELER DİZİNİ.....   | xi   |
| SİMGELER VE KISALTMALAR .....  | xii  |
| 1. GİRİŞ.....  | 1    |
| 2. VERİ AMBARI .....   | 4    |
| 2.1. Veri Ambarı Bileşenleri.....  | 4    |
| 2.1.1. Verinin Nezaman ve Nasıl Toplanacağı .....                        | 4    |
| 2.1.2. Hangi Şema Kullanılacak .....                                     | 5    |
| 2.1.3. Veri Temizliği .....  | 5    |
| 2.1.4. Düzeltmeler Nasıl Yayılacak.....                                  | 5    |
| 2.1.5. Hangi Bilgi Özetlenecek .....                                     | 5    |
| 2.2. Veri Ambarının Tasarlanması .....                                   | 5    |
| 2.2.1. Verinin Çekilmesi.....  | 9    |
| 2.2.2. Veri Dönüştürme.....  | 10   |
| 2.2.3. Veri Yükleme .....  | 12   |
| 2.2.4. Güncelleme .....  | 12   |
| 2.3. İşletimsel Sistemlerle Veri Ambarlarının Farkları .....             | 13   |
| 2.4. Veri Ambarının İşletimsel Sistemden Ayrılması .....                 | 15   |
| 2.5. ÇİAİ Nedir? .....   | 16   |
| 2.6. ÇİAİ Sunucu Türleri.....  | 18   |
| 2.7. Veri Ambarlarından Çok Boyutlu Verinin Elde Edilmesi .....          | 20   |
| 2.7.1. Olaylar .....   | 21   |
| 2.7.2. Boyutlar .....  | 21   |
| 2.7.3. Ölçüm Değerleri.....  | 24   |
| 2.8. Veritabanı Tasarım Metodolojileri.....                              | 27   |
| 2.8.1. Yıldız Şema .....   | 28   |
| 2.8.2. Kar Yağışı Şeması.....  | 29   |
| 2.9. Çok Boyutlu İfadeler – ÇBİ (Multidimensional Expressions-MDX) ..... | 30   |

|   |    |
|---|----|
| 2.9.1. ÇBİ Sözdizim Yapısı .....  | 31 |
| 2.9.1.1 Genel Sözdizim Yapısı .....   | 31 |
| 2.9.1.2 Küme ve Çoklu İşlemleri .....   | 32 |
| 2.9.1.3 Kartezyen Çarpım (CrossJoin) .....  | 34 |
| 2.9.1.4 Filtreleme (Filter) .....   | 34 |
| 2.9.1.5 Sıralama (Order) .....  | 35 |
| 2.9.1.6 Hesaplanan Öğeler (Calculated Members) .....  | 35 |
| 3. VERİ MADENCİLİĞİ .....   | 38 |
| 3.1. Veri Madenciliği Yöntembilimi .....  | 39 |
| 3.2. Veri Madenciliği Modelleri .....   | 40 |
| 3.2.1. Sınıflandırma .....  | 40 |
| 3.2.1.1. Karar Ağacı Algoritması .....  | 42 |
| 3.2.2. Gerileme .....   | 47 |
| 3.2.3. Nitelik Önemi .....  | 47 |
| 3.2.4. Anormallikleri Bulma .....   | 47 |
| 3.2.5. Kümeleme .....   | 48 |
| 3.2.6. İlişkilendirme .....   | 48 |
| 3.2.7. Özellik Çıkarımı .....   | 49 |
| 3.3. Veri Madenciliği Veri Depoları .....   | 49 |
| 3.3.1. İlişkisel Veri Ambarı .....  | 49 |
| 3.3.1.1. Esneklik .....   | 50 |
| 3.3.1.2. Standartlık .....  | 50 |
| 3.3.1.3. En Alt Seviyelerdeki Veri Taneciklerine Kolay Erişim .....   | 50 |
| 3.3.2. ÇİAİ Küpleri .....   | 51 |
| 3.3.2.1. Hız .....  | 51 |
| 3.3.2.2. Kolay Dolaşım .....  | 51 |
| 3.3.2.3. Yüksek Yapısalılıkta Boyutlar .....  | 52 |
| 3.3.2.4. Madencilik Sonuçlarının Küp Yapılarında Saklanması .....   | 52 |
| 4. Veri Ambarı ve OLAP Teknolojilerinden Yararlanılarak Karar Destek Amaçlı Raporlama Aracı Gerçekleştirimi ..... | 53 |
| 4.1. Tez Kapsamında Kullanılan Ürünler .....  | 53 |
| 4.1.1. Oracle 9.2.01. VTYS .....  | 53 |
| 4.1.2. Microsoft SQL Server 2000 VTYS .....   | 53 |
| 4.1.2.1. DTS (Data Transformation Service), Veri Dönüştürme Servisi .....   | 53 |



|   |     |
|---|-----|
| 4.1.3. Microsoft Analysis Services.....   | 55  |
| 4.1.4. Microsoft Visual Studio .NET 2003.....   | 57  |
| 4.2. Veri Ambarı ve Veri Küplerinin Tasarımı ve Gerçekleştirimi.....  | 57  |
| 4.2.1. Veri ambarını oluşturacak ÇİHİ tabanlı ilişkisel veritabanının belirlenmesi.....   | 58  |
| 4.2.2. ÇİHİ sisteminin algılanması ve bu sistem üzerinde ihtiyaç duyulabilecek veri ambarı üzerinde bulunması gereken olayların belirlenmesi..... | 58  |
| 4.2.3. Belirlenen olaylar için veri ambarı olarak belirlenmiş ilişkisel veritabanı üzerinde gerekli tablo tanımlarının yapılması.....             | 59  |
| 4.2.4. Belirlenen olayların beslendiği verilerin temizlenmesi, yeniden biçimlendirilmesi ve veri ambarında uygun tablolara yüklenmesi.....        | 66  |
| 4.2.5. Oluşturulan veri ambarı üzerinde çok boyutlu veri yapılarının tanımlanması.....  | 69  |
| 4.2.5.1. Boyutların uygun sıradüzensel yapıda oluşturulması.....  | 69  |
| 4.2.5.2. Küp Tanımlarının Yapılması.....  | 73  |
| 4.2.5.3. Küp yapılarının özetlenmiş ve detay verileri tutacak şekilde işletilmesi.....  | 75  |
| 4.2.5.4. Hazırlanan küpler üzerinde küp editörü üzerinden sorgulama yapılarak sistemin test edilmesi.....   | 75  |
| 4.2.6. Veri ambarının güncelliğinin korunmasının sağlanması.....  | 78  |
| 4.3. Karar Ağacı Algoritması İle Veri Madenciliği Modelinin Tanımlanması ...  | 80  |
| 4.3.1. İlişkisel Veri Ambarı Üzerinde Karar Ağacı Algoritması.....  | 82  |
| 4.3.2. ÇİAİ Küpleri Üzerinde Karar Ağacı Algoritması.....   | 87  |
| 4.3.2.1. Hasta Bilgileri ÇİAİ Küpü.....   | 87  |
| 4.3.2.2. Hasta Bilgileri ÇİAİ Küpü İle Karar Ağacı Algoritmasının Uygulanması.....  | 89  |
| 4.4. Karar Destek Amaçlı Sorgulama ve Raporlama Aracı Gerçekleştirimi...91  |     |
| 4.4.1. Veri Ambarı ve ÇİAİ Küplerine Erişim Bağlantısının Sağlanması.....   | 92  |
| 4.4.2. Küp Seçimi.....  | 93  |
| 4.4.3. ÇBİ Cümleciğinin Üretilmesi.....   | 94  |
| 4.4.3.1. Boyut – Seviye ya da Yardımcı Öğe Seçimi.....  | 94  |
| 4.4.3.2. İşlem Seçimi.....  | 97  |
| 4.4.3.3. Grid Bileşeni Üzerinde Sorgunun Çalıştırılması.....  | 106 |
| 4.4.3.4. Sorguların Grafikselleşmeleri.....   | 110 |

|   |     |
|---|-----|
| 4.4.4. 'Pivot Table' Servisi .....                                  | 114 |
| 4.4.5. Küp Editörü.....   | 115 |
| 4.4.6. Sorgulardan Rapor Tanımlama .....                            | 116 |
| 4.4.7. Rapor Sonucunun Microsoft Excel Ürünü İle Gösterilmesi ..... | 118 |
| 4.4.8. Karar Ağacı Madencilik Modelinin Çizimsel Gösterimi.....     | 119 |
| 4.4.9. Madencilik Modelinin Sanal Küplerle Sorgulanması .....       | 120 |
| 5. Veri Tabanı ve Veri Ambarı Karşılaştırması .....                 | 122 |
| 6. SONUÇ .....  | 124 |
| KAYNAKLAR.....  | 126 |

## ŞEKİLLER DİZİNİ

|   |    |
|---|----|
| Şekil 2.1 Veri Ambarı Mimarisi .....  | 7  |
| Şekil 2.2 Veri Ambarı Mimarisi .....  | 9  |
| Şekil 2.3 ROLAP Mimarisi .....  | 19 |
| Şekil 2.4 MOLAP Mimarisi.....   | 19 |
| Şekil 2.5 HOLAP Mimarisi .....  | 20 |
| Şekil 2.6 Boyut sıradüzen gösterimi .....   | 22 |
| Şekil 2.7 Küp – Üye Tanımı Akışı.....   | 23 |
| Şekil 2.8 Muayene Küpü Gösterimi .....  | 24 |
| Şekil 2.9 ÇİAİ İşlemleri Çizimsel Gösterimi.....  | 26 |
| Şekil 2.10 Muayene İşlemini Gösteren Varlık-Bağıntı Çizeneği.....   | 27 |
| Şekil 2.11 Varlık-Bağıntı Çizeneğinin Çok Boyutlu Modele Dönüştürümü .....                                | 28 |
| Şekil 2.12 Ameliyat Küpüne Ait Yıldız Şema Gösterimi .....  | 29 |
| Şekil 2.13 Muayene Küpüne Ait Kar Yağışı Gösterimi .....  | 30 |
| Şekil 2.14 Kartezyen Çarpım Grid Örneği .....   | 34 |
| Şekil 2.15 Filtreleme Grid Örneği .....   | 35 |
| Şekil 2.16 Hesaplanan Öğe Grid Örneği .....   | 36 |
| Şekil 2.17 Hesaplanan Öğe Grid Örneği .....   | 37 |
| Şekil 3.1 Genel Karar Ağacı Yapısı Görünümü .....   | 43 |
| Şekil 3.2 Sörf Örneği İçin Oluşan Karar Ağacı.....  | 46 |
| Şekil 4.1 Veri Dönüştürüm Servisi Görev Tanım Arayüzü .....   | 55 |
| Şekil 4.2 Analysis Services Sunucu Mimarisi .....   | 56 |
| Şekil 4.3 Analysis Services İstemci Mimarisi .....  | 57 |
| Şekil 4.4 Ameliyat ve Acil İşlem olaylarını gösteren Varlık-Bağıntı Çizeneği.....                         | 63 |
| Şekil 4.5 Doğum, Hasta ve Günlük Yatan Hasta olaylarını gösteren Varlık-Bağıntı Çizeneği .....            | 63 |
| Şekil 4.6 Muayene ve Laboratuvar İşlemlerini gösteren Varlık-Bağıntı Çizeneği ..                          | 64 |
| Şekil 4.7 Patoloji İşlemleri ve Ölüm olaylarını gösteren Varlık-Bağıntı Çizeneği ...                      | 64 |
| Şekil 4.8 Radyoloji İşlemleri ve Reçete olaylarını gösteren Varlık-Bağıntı Çizeneği .....                 | 65 |
| Şekil 4.9 Tıbbi İşlemler, Yatışı Yapılan Hasta ve Yatak olaylarını gösteren Varlık-Bağıntı Çizeneği ..... | 65 |
| Şekil 4.10 Veri Dönüştürme Servisi Üzerinde Hazırlanmış Pakete Ait Genel Görünüm .....                    | 68 |

|  |    |
|--|----|
| Şekil 4.11 Ameliyat olay tablosunun yüklenmesinde kullanılan betik dili ile hazırlanmış görev içeriği .....  | 69 |
| Şekil 4.12 Muayene Küpü kapsamında Uzmanlık Dalı satırı ve Tarih sütununa göre muayene sayılarını gösteren küp editörü görünümü .....  | 76 |
| Şekil 4.13 Tarih boyutunun ayrıntı düzeyi artırılması (drill-down) sonucu küp editörü görünümü .....   | 76 |
| Şekil 4.14 Muayene Küpü kapsamında Bölümler satırı ve Tarih sütununa göre muayene sayılarını gösteren küp editörü görünümü .....   | 77 |
| Şekil 4.15 Muayene küpü üzerinde bölüm tiplerine göre gruplama seviyesi artırılması (roll-up) sonucu oluşan küp editörü görünümü .....   | 77 |
| Şekil 4.16 2004 yılında, Beyin Cerrahisi Polikliniğinde yapılmış olan muayene sayılarının uzmanlık dalı ve hasta durumuna göre gruplanması sonucunu gösteren küp editörü görünümü (slice – dice) ..... | 78 |
| Şekil 4.17 Veri Dönüştürüm İş Paketi Bünyesinde Bulunan ve Muayene Küpünün Güncellenmesini Sağlayan Görevlere Ait İş Akışı .....   | 79 |
| Şekil 4.18 İlişkisel Veri Ambarı Tablosu İle ‘Üst Ameliyat Açıklama’ Niteliği İçin Oluşturulan Karar Ağacı .....   | 84 |
| Şekil 4.19 İlişkisel Veri Ambarı Tablosu İle ‘Cinsiyet’ Niteliği İçin Oluşturulan Karar Ağacı .....  | 86 |
| Şekil 4.20 Bağımlılık Ağı Çizeneği.....  | 87 |
| Şekil 4.21 HastaBlg Küp İçeriği .....  | 89 |
| Şekil 4.22 ‘Ust Ameliyat Açıklama’ Üye Tanımı İle Oluşturulan Karar Ağacı .....  | 90 |
| Şekil 4.23 ‘Cinsiyet’ Üye Tanımı İle Oluşturulan Karar Ağacı .....   | 91 |
| Şekil 4.24 Karar Ağacını Gösteren Sanal Küp İçeriği .....  | 91 |
| Şekil 4.25 Sistemde Yeralan Küplerin Tutulduğu Liste Yapısı.....   | 94 |
| Şekil 4.26 Sistemde Yeralan Küplerin, Boyutların ve Seviyelerin Tutulduğu Liste Yapıları .....   | 94 |
| Şekil 4.27 Yardımcı Öğeler Veri Giriş ve Güncelleme Ekranı .....   | 95 |
| Şekil 4.28 Ameliyat Tür Boyutu İçin Boyut Seviyeleri Listesi .....   | 96 |
| Şekil 4.29 Yardımcı Öğe Listesi .....  | 96 |
| Şekil 4.30 Sütun Kapsamında Birleştirme İşlemi Örneği .....  | 98 |
| Şekil 4.31 Sıradüzensel gösterim kullanılmadan elde edilen sonuç kümesi.....   | 99 |
| Şekil 4.32 Sıradüzensel gösterim ile elde edilen sonuç kümesi .....  | 99 |

|  |     |
|--|-----|
| Şekil 4.33 Sıradüzensel gösterimle elde edilen sonuç kümesi üzerinde bir alt seviyeye erişim ile oluşan sonuç kümesi .....   | 99  |
| Şekil 4.34 Birinci Sorguya Ait Sonuç Kümesi .....  | 101 |
| Şekil 4.35 İkinci Sorguya Ait Sonuç Kümesi .....   | 101 |
| Şekil 4.36 Sütun Birleşimi Yapılmış Sorguya Ait Sonuç Kümesi .....   | 102 |
| Şekil 4.37 Birinci Sorguya Ait Sonuç Kümesi .....  | 103 |
| Şekil 4.38 İkinci Sorguya Ait Sonuç Kümesi .....   | 103 |
| Şekil 4.39 Satır Birleşimi Yapılmış Sorguya Ait Sonuç Kümesi .....   | 103 |
| Şekil 4.40 Satır ve Sütun Birleşimlerinin Birlikte Kullanımı .....   | 104 |
| Şekil 4.41 4 ÇBİ Cümlecüğinden Oluşan Örnek Sorgu Sonucu (A : Birinci Sorgu Sonucu, B : İkinci Sorgu Sonucu, C : Üçüncü Sorgu Sonucu, D : Dördüncü Sorgu Sonucu) ..... | 106 |
| Şekil 4.42 Grid Bileşeni Görünümü .....  | 107 |
| Şekil 4.43 Grid Bileşeni Görünümü .....  | 108 |
| Şekil 4.44 Grid Bileşeni Görünümü .....  | 109 |
| Şekil 4.45 Grid Bileşeni Görünümü .....  | 110 |
| Şekil 4.46 Dogum küpü üzerinde hazırlanmış ÇBİ sorgu sonucu.....   | 111 |
| Şekil 4.47 'Grafikler' Sayfa Görünümü .....  | 111 |
| Şekil 4.48 'Dogum' küpü üzerinde hazırlanmış sorgunun çubuk çizge görünümü .....   | 112 |
| Şekil 4.49 'Dogum' küpü üzerinde hazırlanmış sorgunun çizgi çizge görünümü .   | 113 |
| Şekil 4.50 'Dogum' küpü üzerinde hazırlanmış sorgunun dairesel çizge görünümü .....  | 114 |
| Şekil 4.51 Pivot Table Servis Arayüzü.....   | 115 |
| Şekil 4.52 Muayene Küp İçeriğini Gösteren Küp Editörü Arayüzü.....   | 116 |
| Şekil 4.53 Raporlama İçin Kullanılan Düğmeler .....  | 117 |
| Şekil 4.54 Rapor ve Sorgu İzleme Formu.....  | 117 |
| Şekil 4.55 " 2004 Yılı İstatistikleri" Adlı Raporun Aktif Rapor Olarak Seçilmesi Durumu .....  | 118 |
| Şekil 4.56 Excel üzerinde hazırlanmış rapor çıktısı .....  | 119 |
| Şekil 4.57 'Cinsiyet' Karar Ağacının Görünümü.....   | 120 |
| Şekil 4.58 Sanal küp üzerinde çalıştırılan ÇBİ sonucu .....  | 121 |

## ÇİZELGELER DİZİNİ

|  |     |
|--|-----|
| Çizelge 2.1 İşletimsel Sistemlerle Veri Ambarının Karşılaştırılması .....        | 13  |
| Çizelge 3.1 Sınıflandırma Algoritmaları Karşılaştırması .....                    | 42  |
| Çizelge 3.2 Sörf örneği için kural oluşturmak için kullanılacak nitelikler ..... | 44  |
| Çizelge 3.3 Sörf Örneği İçin Öğrenme Verileri.....                               | 44  |
| Çizelge 4.1 Veri Ambarında Bulunan Boyut Tabloları ve Görevleri .....            | 59  |
| Çizelge 4.2 Veri Ambarında Bulunan Olay Tabloları ve Görevleri .....             | 61  |
| Çizelge 4.3 Boyut Tanımları .....  | 70  |
| Çizelge 4.4 Küp Tanımları .....  | 73  |
| Çizelge 4.5 Küp Açıklamaları .....   | 74  |
| Çizelge 4.6 Hasta Bilgileri Tablo Yapısı .....                                   | 81  |
| Çizelge 4.7 Hasta Bilgileri İçin Kural Oluşturulabilecek Nitelikler.....         | 81  |
| Çizelge 4.8 Veri Madenciliği Modelinde Kullanılan Nitelikler ve Türleri .....    | 83  |
| Çizelge 4.9 Üst Ameliyat Açıklama' niteliği üretilen karar ağacı kuralları ..... | 83  |
| Çizelge 4.10 'Cinsiyet' niteliği üretilen karar ağacı kuralları örnekleri .....  | 85  |
| Çizelge 4.11 HastaBİg Küpünde Yeralan Boyut Tanımları .....                      | 88  |
| Çizelge 4.12 Veri Madenciliği Modelinde Kullanılan Boyutlar ve Üye Tanımları...  | 90  |
| Çizelge 5.1 Raporu oluşturan sorgu içerikleri .....                              | 122 |
| Çizelge 5.2 Veri Tabanı ve Veri Ambarı Karşılaştırması .....                     | 123 |

## SİMGELER VE KISALTMALAR

|                 |   |
|-----------------|---|
| Bottom –up      | Aşağıdan Yukarıya                               |
| ÇİAİ            | Çevrimiçi Analitik İşleme                       |
| ÇBİ             | Çok Boyutlu İfade                               |
| ÇİHİ            | Çevrimiçi Hareket İşleme                        |
| Data Warehouse  | Veri Ambarı                                     |
| Data Mart       | Veri Pazarı                                     |
| Data Mining     | Veri Madenciliği                                |
| Dimension       | Boyut   |
| Dimension Table | Boyut Tablosu                                   |
| Fact Table      | Olay Tablosu                                    |
| GB              | GigaByte  |
| HOLAP           | Hybrid OLAP – Hibrit ÇİAİ                       |
| İVTYS           | İlişkisel Veri Tabanı Yönetim Sistemi           |
| JDBC            | Java Database Connection                        |
| Measure         | Ölçüm Değeri                                    |
| MDX             | MultiDimensional Expression – Çok Boyutlu İfade |
| MOLAP           | Multidimensional OLAP – Çok boyutlu ÇİAİ        |
| ODBC            | Open Database Connection                        |
| OLAP            | Online Analytical Processing                    |
| OLEDDB          | Open Linking And Embedding for Databases        |
| OLTP            | Online Transaction Processing                   |
| ROLAP           | Relational OLAP – İlişkisel ÇİAİ                |
| SQL             | Structured Query Language                       |
| SVM             | Support Vector Machine                          |
| TB              | TeraByte  |

Top-down

Yukarıdan Aşağıya

VTYS

Veri Tabanı Yönetim Sistemi

YSD

Yapısal Sorgu Dili



## 1. GİRİŞ

Günümüzde büyük ölçüde hareket işleme tabanlı sistemler kullanılır ve şirketler bu sistemler tarafından üretilen geniş sayıdaki bilgiye erişirler. Örneğin, şirket veri tabanlarında müşteriler ve hareketlere ait geniş sayıda nitelikli bilgiler saklanır. Saklanan bilginin boyutu gigabyte hatta terabyte düzeyine erişir. Bu tarz büyük veritabanları, hangi ürün stoklanmalı, ne kadar indirimle gidilmeli gibi ticari kararlarının alınmasında kullanılır. Örneğin, bir hastane yönetimi son bir yıl içerisinde en sık karşılaşılan vakaları ve bu vakalar kapsamındaki ortak özellikleri elde edebilir. Bu bilgi ışığında hastane araştırma-geliştirme biriminin hedefi, bu vakalara çözüm bulmak ve oluşumunu sağlayan etmenleri tespit etmek amaçlı strateji geliştirmek olabilir.

Karar destek kapsamında verinin saklanması ve erişimi bir çok önemli yaklaşımın oluşmasını sağlar (Korth H.F. et al,1991).

- Bir çok karar destek sorgusu YSD ile yazılabileceği gibi, bir çoğunun YSD ile ifade edilmesi imkânsızdır ya da çok zordur. Veri tabanının çok büyük sığalara ulaşmasına rağmen, ÇİAİ veri analizi için kullandığı teknik ve araçlarla, sorgulara anlık cevaplar verebilmektedir.
- Veritabanı sorgu dilleri detaylı istatistiksel analiz için uygun performansta değildir. Veri ambarı ve ÇİAİ uygulamaları için SAS, SAP ve S++ gibi çeşitli yazılım paketleri vardır. Bu paketler büyük ölçekli verilerin veri tabanından alınmasını sağlarlar.
- Bilgi-buluş teknikleri otomatik olarak veriden istatistiksel kuralların ve örüntülerin elde edilmesini sağlar. Veri madenciliği yapay us araştırmacıları ve istatistik çözümleyiciler tarafından bulunan bilgi buluş tekniklerinin birleştirilmesidir.
- Büyük firmaların ticari kararları alırken kullanmak durumunda oldukları çok sayıda veri kaynağı vardır. Veri kaynakları verileri farklı şemalar altında saklarlar. Farklı veri kaynaklarına erişmesi gereken sorguları çalıştırmak için veri ambarları kurulur.

Veri ambarı ve ÇİAİ karar destek sistemlerinin vazgeçilmez elemanları durumuna gelmiştir. Karar destek sistemleri, geleneksel ÇİHİ'ne göre farklı ihtiyaçlar göstermektedir. Veri ambarı, bilgi işleyicilerine (yönetici, idareci, analist) daha iyi ve hızlı karar alımını sağlayan karar destek teknolojileri koleksiyonudur. (Chaudhuri S, et al., 1996) Veri ambarı teknolojisi, üretim, satış, finans servisleri (risk ve kredi kartı analizi gibi), taşımacılık, telekomünikasyon, sağlık gibi birçok sektörde başarıyla uygulanmaktadır. Ralph Kimball'a göre veri ambarı, özellikle analiz ve sorgulama amaçlı yapılandırılmış hareket verilerinin kopyasıdır. (Kimball R.,2002) W.H. Inmon'un tanımıyla veri ambarı, temel olarak organizasyonel karar alımında kullanılan konu yönlendirmeli, bütünlük, zaman değişimli, kararlı/değişime uğramayan veri koleksiyonudur (Chen Zl.,2001). Bu başlıkları biraz daha yakından inceleyelim;

Konu Yönlendirmeli; bir veri ambarı, tüketici, tedarikçi firma, ürün ve satış gibi önemli özneler etrafında kurulur. Veri ambarı bir organizasyonun günlük işleri ve hareket işleme faaliyetleri üzerinde yoğunlaşmak yerine karar verecek kimseler için veriye ait modelleme ve analiz üzerinde yoğunlaşır. Bundan dolayı, veri ambarları tipik olarak ayrıntılı konular çevresinde, karar destek sürecinde faydalı olmayan veriyi dışarıda tutarak basit ve öz bir bakış sağlar.

Bütünlük; bir veri ambarı genellikle ÇİHİ ile üretilmiş ilişkisel veritabanlarında tutulan kayıtlar ile çeşitli farklı türde (heterojen) dosyalarda yer alan verileri bütünlük oluşturur. Veri temizleme ve veri tümeleme teknikleri, isimlendirmede, şifreleme yapılarında, nitelik ölçütlerinde ve benzeri konularda tutarlılığı garantilemek için uygulanır.

Zaman Değişimli; veriler tarihi bir bakış açısından bilgi sağlamak için depolanır. Veri ambarı içerisinde her anahtar yapı, zamanın bir elemanı olarak kesinlik içerir.

Kararlı; bir veri ambarı her zaman, işletimsel çevre içerisinde bulunan uygulama verisinden dönüştürülen verilerin fiziksel olarak ayrı bir kayıttır. Bu ayrıma bağlı olarak bir veri ambarı hareket işlemeyi, geri almayı ve rastlantısal kontrol mekanizmalarını gerektirmez. Veriye erişim için çoğunlukla sadece iki işlem gerektirir: verinin ilk yüklemesi ve veriye erişim.

Genel olarak veri ambarı organizasyonu işletimsel veri tabanından bağımsızdır. Bunun için bir çok sebep vardır. Veri ambarı ÇİAİ işlemlerini destekler ve işlevsellik ve performans gereksinimleri yönünden ÇİHİ işlemlerinden çok farklıdır. Tarihsel, özet ve birleştirilmiş veri detaylı ve bireysel kayıtlardan çok daha önemlidir. Veri ambarında işyükü, yoğun ve karmaşık sorgulardır ve bu sorgular için milyonlarca kayıt üzerinde bir çok tarama, birleştirme ve kümeleme işlemleri yapılır. Sorgu sonuçları ve cevap verme zamanı hareket sonuçlarından çok daha önemlidir.

Karar destek sistemleri işletimsel veritabanlarında unutulmuş veriye ihtiyaç duyar. ÇİHİ işlemleri anlık veriyi içerirken ÇİAİ işlemlerinde tarihsel veriler kullanılır. Karar destek sistemleri çoğunlukla farklı heterojen veri kaynaklarından alınan birleştirilmiş verileri içerir. Farklı veri kaynakları değişik kalitede ve tutarsız verinin elde edilmesine sebep olur. Sonuç olarak, çok boyutlu veri modelleri ve işlemlerini içeren tipik ÇİAİ sistemleri, özel veri organizasyonlarına, erişim ve uygulama metotlarına ihtiyaç duyarlar. Bu ihtiyaçlar ÇİHİ amaçlı kurulmuş veri tabanı sistemlerince karşılanmaz.

## **2. VERİ AMBARI**

Büyük organizasyonlarda karışık organizasyon yapıları ve farklı veriler farklı yerleşim birimlerine dağılmış durumda olabilir. Karar destek kapsamında tüm veri kaynaklarına erişmek gerekecektir.

Veri ambarları, birçok farklı kaynaktan elde edilen bilgilerin depolandığı arşiv alanlarıdır. Tahmin etme ve karar alma işlemlerinin desteklendiği kullanımlarda, veri ambarı veri organizasyonları için merkezi bir depo rolündedir. Organizasyonun ayrıntılı ve homojen görüntüsünü oluşturur.

Geleneksel veritabanı sistemleri, kullanıcı hareketlerine bağlı günlük işlemleri desteklemek için tasarlanmıştır ve bu sistemler işletimsel ya da hareketsel (operational / transactional) sistemler olarak adlandırılır. İşletimsel sistemler hareket ya da işlem yönlendirmeli, veri ambarları ise konu yönlendirmelidir. Örneğin, yüksek seviyeli ticari verinin esnek analitik işlenmesini sağlar. Veri ambarı, veritabanı hareketinden çok sorgulama ve analiz için kullanılmak üzere tasarlanmış ilişkisel bir veritabanıdır. Genelde hareket verisinden elde edilmiş tarihsel bilgileri içerdiği gibi başka dış kaynaklardan gelen verileri de içerebilir. Veri tabanı hareketlerinden kaynaklanan iş yüküyle analiz yükünü birbirinden ayırır. Bu sayede değişik kaynaklardan toplanan verilerin daha kolay bir şekilde organize edilmesini sağlar.

Veri ambarının en önemli özelliği yüklenen verinin salt okunur ve değişken olmayan yapıda olmasıdır. İşletimsel veri genellikle gerçek zamanlıdır fakat veri ambarında tarihsel veriler mevcuttur. Veri ambarında yüksek seviyeli tarihsel veri genellikle gelecekte ne olabileceğini tahmin etmek için raporlama ve analiz amacıyla kullanılır.

### **2.1. Veri Ambarı Bileşenleri**

#### **2.1.1. Verinin Nezaman ve Nasıl Toplanacağı**

Verilerin toplanması kapsamında kaynak tabanlı ve hedef tabanlı olmak üzere iki tür mimariden söz edebiliriz (Jarke M. et al.,2000). Kaynak tabanlı mimaride, veri kaynakları yeni bilgiyi iletir ya da periyodik olarak bilgi iletimi gerçekleştirir. Veri ambarını besleyen veri kaynakları her veri değişiminde ya da belirlenmiş bir

zaman aralığında güncel verileri veri ambarına aktarırlar. Hedef tabanlı mimaride veri ambarı periyodik olarak kaynaklardan yeni verileri alabilmek için istemde bulunur. Verilerin ne zaman toplanması gerektiği, veri ambarını oluşturan sistemin güncel veri gereksinimine olan ihtiyacına göre belirlenir.

### **2.1.2. Hangi Şema Kullanılacak**

Veri kaynakları aynı kapsamdaki verileri farklı şemalar ve veri yapıları altında saklayabilir. Bu durumda veri ambarı şema entegrasyonu yaparak verilerin entegre bir şema üzerine dönüştürülmesini sağlar. Dolayısıyla veri ambarı, veri kaynakları ile tıpatıp aynı veri yapısına sahip değildir.

### **2.1.3. Veri Temizliği**

Verinin doğrulanması ve ön işleme veri temizliği '*data cleansing*' olarak adlandırılır. Veri kaynakları bir çok veri tutarsızlığı ile verilerini gönderirler. Örneğin mahalle, cadde isimleri yazım yanlışına sahip olabilir ya da posta kodları hatalı olabilir. Bu tür durumlarda, her şehir için mahalle, cadde, sokak isimleri ve posta kodları saklanarak gerekli hatalar çapraz kontrollerle düzeltilebilir.

### **2.1.4. Düzeltmeler Nasıl Yayılacak**

Veri kaynaklarında yapılan veri değişiklikleri veri ambarına yansıtılmalıdır. Eğer kaynaktan ve hedefte yer alan veri modeli aynı ise doğrusal bir veri aktarımı ile sorun çözülecektir. Veri ambarındaki veri modeli kaynaktan farklıysa 'görüntü bakımı-view maintenance' kapsamında güncelleme gerçekleştirilir.

### **2.1.5. Hangi Bilgi Özetlenecek**

Hareket-işleme sistemleri ile elde edilen veri büyük kapasitelere ulaşacaktır. Bazı veriler için, verinin kendisi yerine özetlenmiş şeklinin saklanması daha doğru olabilir. Örneğin, mağaza sisteminde her bir satışın saklanması yerine isim ve kategori kapsamında toplam satışlar saklanabilir.

## **2.2. Veri Ambarının Tasarlanması**

Veri ambarının kurumu ve bakımı kapsamında çeşitli mimariler mevcuttur. Birçok organizasyon, kurumun tümünü içeren bütünleşik ambar oluşumunu ister, fakat bu yöntem çok uzun ve karmaşık bir sürece sahiptir ve maliyetlidir. Bu yöntem yukarıdan aşağıya (top-down) doğru tasarım olarak adlandırılır. Bu yaklaşım

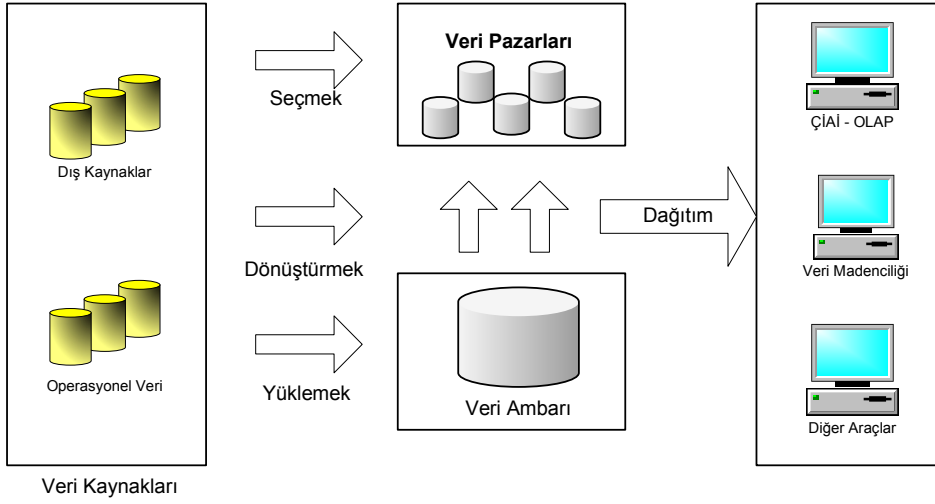
ayrıntılı bir tasarım ve planlama ile başlar. Teknolojinin gelişmiş olduğu ve iyi bilindiği durumlarda ve mutlaka çözülmesi gereken iş problemlerinin açık ve iyi anlaşıldığı durumlarda kullanışlıdır. Bir veri ambarının yukarıdan aşağıya gelişimi, sistematik bir çözüm olarak hizmet eder ve bütünleştirme problemlerini minimize eder. Bunun yanı sıra pahalıdır, geliştirilmesi uzun zaman alır. Tüm organizasyon için bilinen bir veri modelinde anlaşma ve veri modelini oluşturmaya ilişkin tutarlılıkta esneklik yönünden zayıftır. Bu yönüde alternatif olarak veri pazarlarının (data mart) oluşturulmasını sağlayan aşağıdan yukarıya (bottom-up) doğru tasarım yöntemi ortaya çıkmıştır. Aşağıdan yukarıya doğru tasarım yöntemiyle organizasyonun tümü yerine organizasyon içerisindeki alt bölümler için küçük veri ambarlarının her biri veri pazarlarını oluşturur. (Kimball R., 2002) Bu yaklaşım deneyler ve prototiplerle başlar. İş modellemenin ve teknoloji gelişiminin erken aşamalarında kullanışlıdır. Aşağıdan yukarıya doğru yaklaşım esneklik, düşük maliyet ve yatırımların kısa zamanda geri dönüşünü sağlar. Her iki yöntemin de kullanıldığı birleşik yaklaşımda, organizasyon aşağıdan yukarıya doğru yaklaşımın hızlı uygulaması ve fırsatlar sunan kullanımına sahipken, yukarıdan aşağıya doğru yaklaşımın planlı ve stratejik özelliğini de kendi yararına kullanabilir.

Veri ambarı mimarisi, birçok işletimsel veri tabanından ve dış kaynaklardan verinin alınmasını, verinin temizlenmesini, dönüştürülmesini ve bütünleştirilmesini, verinin veri ambarına yüklenmesini ve güncellemelerin periyodik olarak veri ambarına yansıtılabilmesini sağlayan araçları içerir. Ana veri ambarının yanında bir çok veri pazarı bulunabilir. Veri ambarı ve veri pazarlarında saklanan veri bir ya da birçok ambar sunucusu tarafından yönetilir, çok boyutlu veri rapor, sorgu ve analiz araçları ile veri madenciliği araçları için dağıtılır. Son olarak, metaveri'nin saklanması ve yönetimi için bir depo ve ambar sisteminin izlenmesi ve yönetimi için araçlar yer alır.

Veri ambarı birçok bilgi kaynağına erişebilmelidir;

- Veritabanı sistemleri (ilişkisel, nesneye yönelik, ağ,..)
- Dış bilgi kaynakları (diğer kurumlardan elde edilen bilgiler)
- Standart uygulama dosyaları (Excel, Cobol ...)

- Diğer dokümanlar (Word, www, ...)



Şekil 2.1 Veri Ambarı Mimarisi

Veri ambarı mimarisini 3 katmanlı olarak tanımlayabiliriz.

1.Katman; ilişkisel veritabanı sistemi olan veri ambarı veritabanı sunucusudur. Birden çok dış kaynaktan gelen veriler temizlenip güvenli hale dönüştürülerek bu sunucu üzerinde depolanır. Veri ambarları çok çeşitli veri çekme ve temizleme araçları ile yükleme ve güncelleme paketleri kullanır. Dış kaynaktan veri çekmek için standartlaşmış arayüzler mevcuttur. ODBC, OLEDB, JDBC, Oracle Open Connect, Sybase Enterprise Connect, Informix Enterprise Gateway örnek olarak verilebilir. Bu arayüzler VTYS tarafından desteklenmektedir ve istemciler üzerinde yürütülecek YSD sorgularının çalıştırılmasını sağlarlar.

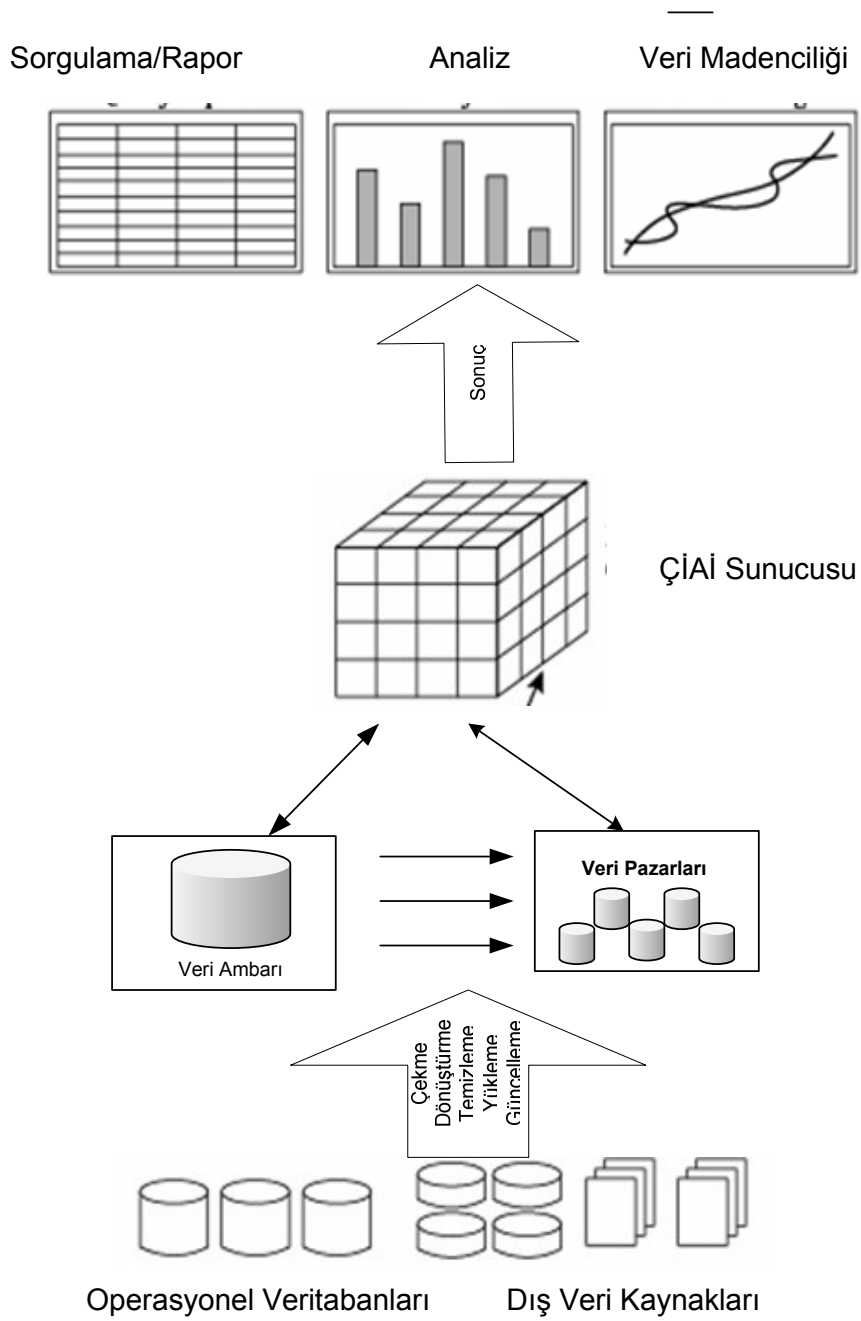
2.Katman; ilişkisel veritabanında tutulan temizlenmiş veriler üzerinden çok boyutlu verilerin üretildiği katmandır. Bu katman ilişkisel veritabanı üzerinde olabileceği gibi (ROLAP – İlişkisel ÇİAI yöntemi), veritabanından tamamen bağımsız (MOLAP – Çok boyutlu ÇİAI yöntemi) ya da veritabanına kısmen bağımlı (HOLAP – Hibrit ÇİAI yöntemi) olarak da gerçekleştirilebilir.

3.Katman; Sorgu ve raporlama araçları, analiz araçları, veri madenciliği araçları gibi karar destek kapsamında uç kullanıcıyla bağlantılı katmandır.

Veri ambarının tasarlanması karmaşık bir iştir ve aşığıdaki adımların takip edilmesi gerekir; (Kimball R.,2002)

- Mimariyi belirle, kapasite planlaması yap, sunucu, veritabanı, ÇİHİ sunucu ve araçları seç
- Sunucular, saklama birimleri ve istemci araçları entegre et
- Veri ambarı şema ve görüntülerini tasarla
- Fiziksel ambar organizasyonunu tanımla (veri yerleşimi, bölümlenme ve erişim metotları)
- Veri kaynaklarına bağlan
- Veri çekmek, temizlemek, taşımak, yüklemek ve güncellemek için betikler hazırla
- Verileri çek
- Verileri temizle (kaynak veri üzerindeki tutarsızlıkların temizlenmesi)
- Veri dönüştürmelerini yap (farklı veri formatlarına ya da dillere dönüştürüm)
- Yüksek hızlı veri transferi imkânını sağla
- Verileri yükle (Verilerin veri ambarına yüklenmesi)
- Şema, görüntü, betik ve diğer metaveri bilgileri ile depoyu oluştur
- Veri kalitesinin kontrol edilmesi (doğruluk ve bütünlük açısından)
- Son kullanıcı uygulamalarını tasarla ve gerçekleştir
- Ambar ve uygulamaları bir araya getir





Şekil 2.2 Veri Ambarı Mimarisi

### 2.2.1. Verinin Çekilmesi

İşletimsel sistemlerde ve dış kaynaklarda saklanmış olan verilerin veri ambarına çekilmesi işlemidir. Bu işlem ile verinin temizlenmesi, veri ambarında tanımlı kısıtlara göre uyarlanması gibi işlemler gerçekleştirilmez. Sadece hangi tür verilerin ambar kapsamına alınacağı belirlenir.

### 2.2.2. Veri Dönüştürme

Veri ambarları karar destek amaçlı kullanıldığı için ambar üzerindeki verinin doğru olması çok önemlidir. Birçok dış kaynaktan alınan büyük ölçekli veri olduğu için hataların ve tutarsızlıkların yer alması normaldir. Bu nedenle hataları temizleyen araçların önemi artmaktadır. Bazı hata örnekleri; tutarsız alan uzunluğu, tutarsız açıklamalar, tutarsız değer atamalar, unutulmuş girişler ve çiğnenmiş tutarlılık kısıtları olarak verilebilir. Üç tür veri temizleme aracından bahsedebiliriz. (Korth H.F. et al,1991) Veri geçirme araçları (data migration tools), bir sözcüğün başka bir sözcükle yer değiştirmesi gibi basit kuralların tanımlanmasını sağlar. Prism tarafından çıkarılan 'Warehouse Manager' bu türdeki en popüler araçtır. (Chaudhuri S, et al., 1996) Veri arama araçları (data scrubbing tools) alana özel bilgiler ışığında işlem yaparlar. Çoğunlukla ayrıştırma ve fuzzy eşleştirme teknikleri ile temizleme yaparlar (posta adreslerinin doğrulanması gibi). Veri denetleme araçları (data auditing tools), verinin taranması için kural ve ilişkiler tanımlar. Veri madenciliği araçları bu bağlamda düşünülebilir.

Veri dönüştürme, basit dönüştürme, temizleme, kaynaştırma, kümeleme ve özetleme şeklinde dört gruba ayrılabilir. (Bischoff J. Et al. 1997)

- **Basit dönüştürme**, belirli bir zaman diliminde bir veri içeriğinin, onun bağlı bulunduğu diğer bilgilerle olan ilişkisi önemsenmeden değiştirilmesidir. En yaygın kullanılan basit dönüştürme türü, veri türü dönüştürmesidir. Veri kaynağında bulunan veri türünün veri ambarına aktarılırken veri türünün, uzunluğunun, bütünlük kısıtlarının değiştirilmesini sağlar. Örneğin boş bırakılabilir olarak belirlenmiş bir alanın veri ambarında boş olamaz kısıtına sahip olması istenebilir. Number(9) olarak tanımlı bir alanın Number(13,2) olarak değiştirilmesi sağlanabilir. Diğer bir basit dönüştürme türü, tarih/zaman formatı dönüştürmesidir. Bir çok işletimsel sistemde farklı tarih formatları kullanılabilir. Bu sistemlerden veri ambarına aktarılacak olan tarih verilerinin aynı tarih formatında olması gerekmektedir. Eski işletimsel sistemler için 2 hane olarak saklanmış yıl bilgisinin 4 haneye çıkartılması gibi dönüştürmeler bu tür altında yer alır. Alan şifrelerinin çözülmesi işlemi, basit dönüştürme altında yer almaktadır. Farklı işletimsel sistemlerde aynı veriye karşılık gelen farklı alan tanımları ve kısıtları yer alabilir. Örneğin, cinsiyet alanı için bir

sistemde 'E','K' içeriđi tutulurken bařka bir sistemde 'ERKEK','KADIN' bilgisi kullanılabilir. Bu tür verilerin dönüřtürmesi için, veri ambarında tanımlanan yapıya göre uygun dönüřtürmeler yapılmalıdır.

• **Temizleme**, veri formatının düzenlenmesinden ziyade alanların gerçek içeriklerinin dönüřtürülmesi işlemidir. En çok kullanılan temizleme işlemi geçerli deđerlerin dönüřtürülmesidir. İşletimsel sistemlerden aktarılacak olan verilerin aralık deđerlerinin kontrol edilmesi, kabul edilebilir deđerler listesine uygunluđunun belirlenmesi, bađımlılık kısıtlarının kontrol edilmesi bu kapsamda yer alır. Verinin 1000-9999 arasında olmasının sađlanması ya da muayene işlemi yapan doktorun doktorlar listesinde var olması gerekliliđinin kontrol edilmesi örnek olarak verilebilir. Karmařık yeniden formatlama yöntemi ile farklı formatlarda tutulan verilerin ambar üzerinde tek biçimli görüntenmesi sađlanır. Özellikle adres içerikli verilerde bu yöntem kullanılır. Çok farklı řekillerde saklanan fakat aynı adrese karřılık gelen içeriklerin tek tipte tutulması sađlanır.

• **Kaynařtırma**, bir ya da daha çok kaynakta saklanan işletimsel verilerin, veri ambarı üzerinde tutulan yeni veri yapısına göre, alanlarının bire bir eřleřtirilmesi işlemidir. Eřleřtirme işlemi bir çok sistemde %80 oranında basit eřleřtirme olarak adlandırılan ve alanların hedef ve kaynak arasında bire bir eřleřtirilmesiyle oluřur. Fakat %20 oranında karmařık eřleřtirme gerekmektedir. Örneđin, birden çok hastaneye ait verilerin tek bir veri ambarında toplandıđını düşünelim. Bir hasta birden çok hastaneden hizmet almıř olabilir ve her işletimsel sistemde farklı bir belirteç ile tanımlanmıř olabilir. İlgili hastaya ait verilerin veri ambarına aktarılmadan önce biricik belirteç altında toplanması gerekmektedir. Bařka bir karmařık eřleřtirme durumu aynı verinin birden çok kaynakta bulunması durumunda ortaya çıkar. Bu durumda hangi verinin en son güncel durumu ifade ettiđi tespit edilerek eřleřtirme ilgili kaynak kullanılarak yapılmalıdır.

• **Kümeleme ve özetleme**, işletimsel sistemlerde yer alan günlük her türlü hareketin tutulduđu detaylı verilerin veri ambarının tanımlanma amacına göre gruplanıp özet veriye dönüřtürülmesi işlemidir. Veri ambarının genel mantıđına baktıđımızda detaylı veri yerine daha genel özet verinin varlıđını

görmekteyiz. Bu durumda veri ambarında saklanan kayıt sayısı da azalmaktadır. Örneğin bir hastane uygulama yazılımında gün içerisinde yapılan her türlü işlemler kaydedilmektedir. Bu verilerden oluşan veri ambarında ise sadece aylık olarak yapılmış işlemlerin türleri ve belirleyici içerikleri olmaktadır. Şubat ayı içerisinde anestezi yapılarak ameliyat edilmiş yatan hasta sayısının veri ambarındaki değeri işletimsel sistemden günlük bazda alınan ameliyat bilgilerinin özetlenmesiyle elde edilir.

### **2.2.3. Veri Yükleme**

Verinin çekilmesi, temizlenmesi ve dönüştürülmesinden sonraki adım verinin ambar üzerine yüklenmesi işlemidir. Bu adım öncesinde de ek işlemler gerekebilir: bütünlük kısıtlarının kontrol edilmesi, sıralama, özetleme, kümeleme, dizin tanımlama, bölümlenme gibi.

Yükleme paketi sistem yöneticisine mevcut durumun izlenmesini sağlayabilmeli, iptal, askıya al ve devam et işlemlerine izin verebilmeli ve hata durumunda veri bütünlüğü bozulmayacak şekilde yeniden işleme başlayabilmeyi sağlamalıdır.

### **2.2.4. Güncelleme**

Ambarın güncellenmesi, kaynak veriler üzerinde gerçekleşen güncellemelerin ilişkili ana veriler üzerine yansıtılmasıdır. Burada iki önemli nokta vardır, güncelleme işlemi ne zaman ve nasıl yapılmalı. Çoğunlukla ambar periyodik olarak güncellenir. Güncelleme politikası ambar yöneticisine, kullanıcı ihtiyaçlarına ve farklı veri kaynaklarına göre değişkenlik gösterir. Güncelleme teknolojileri kaynakların karakteristiklerine ve veritabanı sunucusunun yeteneklerine de bağlıdır. Birçok veritabanı sistemi replikasyon sunucuları kullanır. Kaynak veri değiştiği zaman replikasyon sunucu ambar üzerindeki veriyi de günceller. İki temel replikasyon tekniği vardır, veri taşıma ve hareket taşıma. (Chaudhuri S, et al., 1996)

Veri taşıma yönteminde, ambar üzerindeki tablo kaynak veritabanı üzerindeki tablonun bir kopyası olarak düşünülür. 'After\_row' tetikleyicileri ile kaynak veri tabanında gerçekleşen güncelleme işlemleri hedef veri tabanına yansıtılır. Bu 'Oracle Replication Server' tarafından kullanılan bir yöntemdir.

Hareket taşıma yönteminde ise tetikleyiciler ve özel günlük tabloları yerine, hareket günlükleri kullanılır. Kaynak bölümde hareketler günlük dosyalarına kaydedilir ve bu dosyalar hedef veritabanına transfer edilir. Bu yöntem ile tetikleyicilere gereksinim olmadığı için işletimsel veritabanlarının işyükü azaltılır. Sysbase Replication Server ve Microsoft SQL Server bu yöntemi kullanır.

### 2.3. İşletimsel Sistemlerle Veri Ambarlarının Farkları

Çizelge 2.1 İşletimsel Sistemlerle Veri Ambarının Karşılaştırılması

| <b>İşletimsel Sistemler</b>   | <b>Veri Ambarı Sistemleri</b>   |
|---|---|
| Az seviyede raporlama içeren ve yüksek seviyeli hareket işleme amaçlı tasarlanmış sistemlerdir. (ÇİHİ)  | Analiz ve raporlama amaçlı tasarlanmış ve analitik işlemleri (ÇİAİ) destekleyen sistemlerdir.   |
| Genellikle işlem yönlendirmeli ya da işlem tabanlı (process-oriented / process-driven) sistemlerdir. Belirli ticari işlemleri ya da görevleri yerine getirmek için tasarlanırlar. | Veri ambarı sistemleri konu yönlendirmelidir. (subject-oriented) Konu alanları çoğunlukla bir ya da birden fazla işletimsel sistem verisinden oluşur.                     |
| Genellikle anlık veri ile ilişkilidir.  | Genellikle tarihsel ve özet veri ile ilişkilidir.   |
| İhtiyaç durumuna göre veri genellikle güncellenir.  | Genellikle değişkenlik göstermeyen veriler içerir. Yeni veriler eklenir fakat bir kez yüklenir, çok nadiren veri güncellenir. Dolayısıyla veri çoğunlukla salt okunurdur. |
| Hızlı ekleme ve güncelleme amacıyla eniyileme yapılır.  | Yüksek yoğunluktaki veri üzerinde hızlı veri erişimi amacıyla eniyileme yapılır.  |
| Genellikle uygulamaya özel sistemlerdir. Birbirinden bağımsız uygulamalar tekrarlı veri oluşumunu doğurur.  | Uygulama katmanında sistemler entegre edilir ve veri tekrarı engellenir.  |
| Her kayıt analiz açısından pek çok gereksiz bilgi içerir.   | Analiz açısından her kayıt büyük önem taşır.  |

Çizelge 2.1 İşletimsel Sistemlerle Veri Ambarının Karşılaştırılması (devam)

| <b>İşletimsel Sistemler</b>   | <b>Veri Ambarı Sistemleri</b>   |
|---|---|
| Karmaşık sorguları yaratmak ve işleme koymak zahmetlidir.   | Karmaşık sorgulara gerek kalmadan ihtiyaç duyulan verilere kolayca erişilebilir.                  |
| Karmaşık sorguların sonuçları saatlerle ölçülebilen zaman dilimlerini kaplayabilir.   | Sorgular kısa sürede cevaplanır.  |
| Sisteme girilmiş her türlü veri mevcuttur. Kullanıcı hataları ve eksik bilgiler bulunabilir.                                    | Sadece ayıklanmış ve güvenilir veri kullanılır.   |
| Nihai kullanıcının çok fazla bilgisayar becerisinin bulunması gerekmez.   | Deneyimli kullanıcılara yöneliktir, yüksek seviyeli bilgisayar becerisi gerekir.                  |
| Müşteri merkezlidir, bilgi teknolojisi profesyonelleri ve müşteriler tarafından işlembilgi ve sorgulama işleme için kullanılır. | Pazar merkezlidir ve analistler, uzmanlar ve yöneticiler tarafından veri analizi için kullanılır. |
| Varlık-bağıntı veri modelini kullanır ve uygulama merkezli veritabanı tasarımı vardır.  | Yıldız ya da kar akışı modelini kullanır ve konu yönlendirmeli bir veri tabanı tasarımı vardır.   |
| Veri ekleme, güncelleme ve silme sıklığı çok fazladır. Güncelleme sürekli yapılır.  | Belirli zaman aralıklarında veri ekleme gerçekleştirilir.   |
| Sisteme erişen kullanıcı sayısı fazladır.   | Karar alıcı kademedeki bulunan belirli sayıdaki kullanıcılara yöneliktir.                         |
| Fazla detaylı güncel veriyi yönetir.  | Karar destek amaçlı veri analizi için kullanılır. Büyük miktarlarda tarihsel veriyi yönetir.      |
| Fiziksel kapasitesi MB ya da GB seviyelerindedir.   | Fiziksel kapasitesi GB ya da TB seviyelerindedir.   |

## 2.4. Veri Ambarının İşletimsel Sistemden Ayrılması

İşletimsel sistemin kurulu olduğu bir ilişkisel veri tabanının mevcut olduğu durumlarda, veri ambarı için ayrı bir ilişkisel veritabanına ihtiyaç vardır. Bu tür durumlarda “mevcut veri tabanı üzerinden doğrudan ÇİAİ gerçekleştirimi yapmak yerine neden ayrı bir veri ambarı kurmak için ek zaman ve kaynak kullanmalıyız?” sorusu oluşmaktadır. Veri ambarı ile işletimsel verilerin yer aldığı veritabanının fiziksel olarak ayrılması gerekliliğinin nedenleri aşağıda açıklanmıştır.

- **Her iki sistem için de yüksek performansın elde edilmesini sağlamak.** İşletimsel veritabanı, birincil anahtar üzerinden indeksleme ve hazır sorguları eniyileme gibi bilinen görev ve iş yüklerinden hareketle tasarlanır. Veri ambarı sorguları büyük yoğunlukta verileri kapsayan karmaşık sorguları içerir. Özetlenmiş seviyelerdeki verilerin büyük gruplarının hesaplanması ile ilgilenir ve özel veri organizasyonu, erişim ve çok boyutlu modellere dayanan yaklaşımlar gerektirir. İşletimsel veritabanlarında ÇİAİ sorgularını çalıştırmak, işletimsel görevlerin performansını büyük ölçüde düşürecektir.
- **İş üretiminde olumsuzluğu önlemek.** İşletimsel veritabanı, çoklu hareketlerin uyumlu işlenmesini destekler. Veri uyumluluğunu sağlamak için kilitleme ve günlük kayıt tutma, veri güvenilirliği ve sağlamlığı için kurtarma ve geri yükleme mekanizmaları kullanır. ÇİAİ sorguları büyük çoğunlukla, kümeleme ve özetleme için geniş ölçekli veri kayıtlarına salt okunur erişime gereksinim duyar. Uyumluluk kontrolleri ve kurtarma mekanizmaları ÇİAİ işlemlerine uygulanırsa işlemlerin hızlı ve kolay yürütülmesine engel olur. Dolayısıyla iş üretimi üzerinde olumsuz etkileri oluşur.
- **İşlevsellik farklılıklarının yarattığı bakım ve yönetim sorunlarından kurtulmak.** Karar destek için tarihsel bilgi gerekli iken, işletimsel sistemler genel olarak geçmiş verilerle ilgilenmezler. Bu bağlamda, işletimsel veritabanlarında veri çok olmasına rağmen, karar vermek için gerekli netlikten uzaktır. Karar destek için, heterojen kaynaklardan gelen, yüksek kalitede, temiz ve bütünleşik veriye ihtiyaç vardır. İşletimsel sistemler, detaylı ham veri içerirler. İki sistem farklı işlevsellikler sunduklarından ve

farklı içerikte verilerle ilgilendiklerinden tek bir veritabanı sistemi bünyesinde iki ayrı sistemi yönetmek oldukça karmaşıktır ve zorluklar içerir.

## 2.5. ÇİAİ Nedir?

ÇİAİ terimi, ilk kez Dr. E.F.Codd tarafından 1993 yılında yayınlanan ve “Providing On-Line Analytical Processing to User Analysts – Çözümleyici kullanıcılar için çevrim içi analitik işleminin sağlanması” başlıklı yayın ile ortaya çıkmıştır. (Bischoff J.,et al,1997) Orijinal yayın kapsamında ÇİAİ sistemlerinin 12 kuralından söz edilir, 1995 yılında 6 ek kural daha tanımlanmıştır. OLAP Konseyi'nin verdiği tanıma göre, ÇİAİ, kullanıcı tarafından anlaşılabilir şekilde gerçek boyutlara taşınmış işlenmemiş ham veri üzerinde çeşitli bilgi görüntüleri sunarak, analistler, yöneticiler ve çalışanların veriye hızlı, tutarlı ve etkileşimli biçimde erişmesini sağlayan bir yazılım teknolojisi kategorisidir.

Dr. Codd tarafından tanımlanan 12 kural şu şekildedir; (<http://www.olapreport.com/fasmi.htm>) (Berson A. et al.,1997)

1. **Çok boyutlu Kavramsal Görüntü**, ÇİAİ merkezini oluşturan kuraldır. ÇİAİ sistemi, kullanıcıya çok boyutlu model sağlamalıdır.
2. **Saydamlık**, bir hesap çizelgesi üzerinde ÇİAİ motoru tarafından doldurulmuş sayfalarca veri olabilir fakat verilerin hangi kaynaktan geldiği tam olarak bilinemez. ÇİAİ sistem teknolojisi, alt tarafta kullanılan veritabanı ve hesaplama mimarisi, girdi veri kaynaklarının heterojenliği hakkında kullanıcıya şeffaf olmalıdır. Bu nedenle, ürünler heterojen veri kaynaklarına gerçek iletişim sağlamalıdır.
3. **Erişebilirlik**, ÇİAİ arabulucu olarak kullanılır. ÇİAİ motoru, heterojen veri kaynakları ve ön-uç arasında yer alır. Analizi sağlamak için ihtiyaç duyulan verilere erişebilmelidir.
4. **Tutarlı Raporlama Performansı**, boyutların sayısına ve veritabanının büyüklüğüne bağlı olmaksızın kullanıcı herhangi bir performans farkı hissetmemelidir.



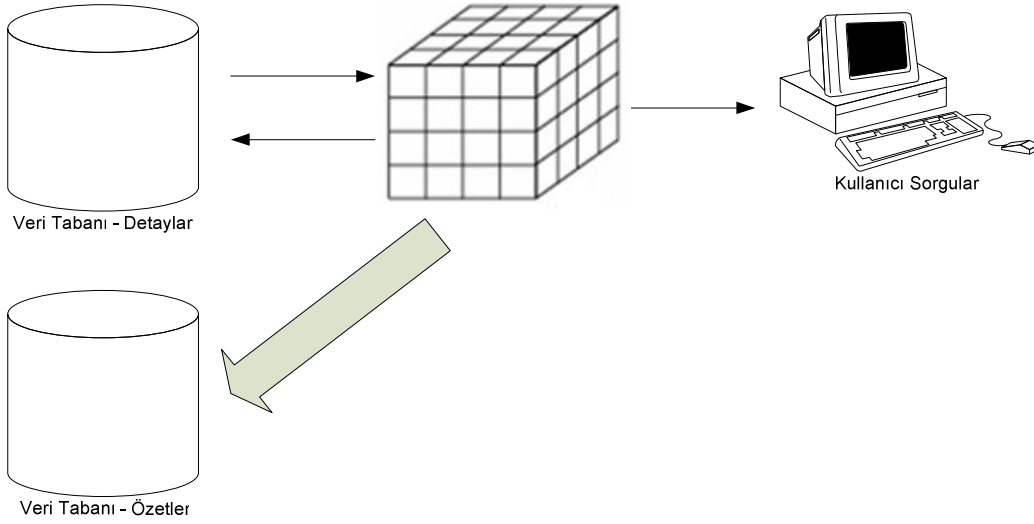
5. **İstemci-Sunucu Mimarisi**, ÇİAİ sistemi istemci-sunucu mimarisinin tüm özelliklerini taşımalıdır. İstemciler en az emek ve programlama entegrasyonu gösteren ÇİAİ sunucudan hizmet alabilmelidir.
6. **Genel Boyutlandırma**, her boyut verisi yapısal ve işletimsel yetenekleri açısından eşit olmalıdır. Temel veri yapıları, formüller ve raporlama formatları herhangi bir boyuta özel olmamalıdır.
7. **Fiziksel Seviyenin Otomatik Ayarlanması**, ÇİAİ sistemi, fiziksel şema görünümünü belirli bir analitik modele, veri hacmine ve veri seyrekliğine göre adapte etmelidir.
8. **Çoklu Kullanıcı Desteği**, ÇİAİ motoru aynı model kapsamında birden çok kullanıcıya eş zamanlı olarak hizmet verebilmelidir.
9. **Çapraz Boyutlama İşlemlerinin Sınırlandırılmaması**, boyut üzerindeki sıradüzensel yapıları desteklemeli ve roll-up işlemlerini otomatik olarak yapabilmelidir. Her türlü hesaplama işlemi tüm boyutlar kapsamında yapılabilmelidir.
10. **Sezgisel Veri İşletimi**, pivot, drill-down, roll-up ve diğer işlemler nokta tıklama, sürükle bırak tarzında işlemler ile kolaylıkla gerçekleştirilebilmelidir.
11. **Esnek Raporlama**, görsel sunumları içerecek analitik raporlama imkânı olmalıdır.
12. **Sınırsız Boyut ve Kümeleme Seviyeleri**, boyut sayısında ve kümeleme seviyelerinde sınır olmamalıdır. Teknik olarak hiçbir ürünün limitleri olan bir bilgisayar üzerinde limitsiz imkânlar sunması mümkün değildir. Bu kuralda belirtilen istenilen sınırların olabildiğince yüksek olmasının sağlanmasıdır.
13. **Toplu Çekme**, ÇİAİ ürünleri, çok boyutlu verilerin fiziksel ortamı için hem veri ambarı kapsamında tuttıkları veri tabanını hem de canlı erişim yaptıkları dış kaynakları önerebilirler. HOLAP sunucu türü bu özelliği kapsamaktadır.

14. **ÇİAİ Analiz Modelleri**, ÇİAİ ürünleri dört analiz modelini de kapsayabilmelidir. Bu modeller, kategorik, açıklanabilir ve yorumlanabilir, derinlemesine düşünebilen ve formüsel (Categorical, Exegetical, Contemplative , Formulaic) olarak tanımlanmıştır.
15. **Normalleştirilmemiş Verinin Uyumu**, ÇİAİ motoru ile normalleştirilmemiş veri kaynaklarının entegrasyonunu içermektedir. ÇİAİ ortamında yapılan veri güncellemeleri, normalleştirilmemiş dış veri kaynaklarına yansıtılmamalıdır.
16. **ÇİAİ Sonuçlarının Kaynak Verilerden Ayrı Olarak Saklanması**, ÇİAİ araçları ile gerçekleştirilen veri okuma, güncelleme ve ekleme işlemleri verilerin gerçek kaynaklarına yansıtılmamalıdır.
17. **Unutulmuş Değerlerin Çekilmesi**, tüm boş değerler tek biçimli bir formata dönüştürülmelidir ve sıfır değeri ile karıştırılmamalıdır.
18. **Unutulmuş Değerlerin Uyumu**, boş bırakılmış değerler ÇİAİ motoru tarafından görmemezliğe gelir. Bu nedenle 17. maddede belirtildiği gibi unutulmuş değerler uygun bir içeriğe dönüştürülmelidir.

## 2.6. ÇİAİ Sunucu Türleri

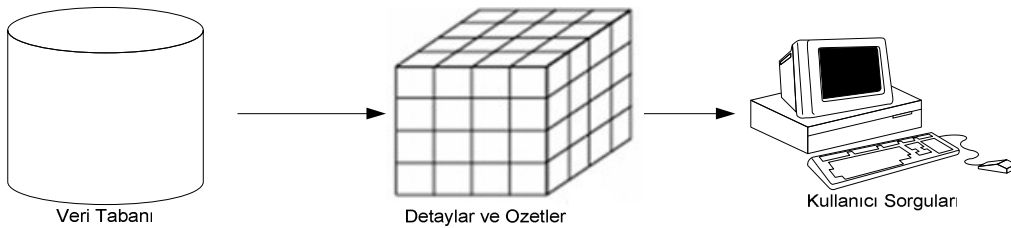
ÇİAİ kapsamında 3 sunucu türü mevcuttur. (Seidman C, 2001)

**ROLAP** (Relational OLAP), ROLAP saklama ünitesi orijinal ilişkisel veritabanını kullanır. VTYS yeni özelliklere sahiptir, veriler yıldız ya da kar akışı şemalar halinde organize edilir. Veri detayları ve özet bilgiler veritabanında tablolar olarak tutulur. Bu nedenle çok boyutlu sorgular için özel mekanizmalara gerek olmadan standart YSD kullanılarak raporlar oluşturulabilir. Çok boyutlu verilerin tablolar ile ifade edilmesi OLAP motoru üzerinden oluşturulan sorguların çalıştırılması oldukça zaman alıcı bir işlem haline dönüşür.



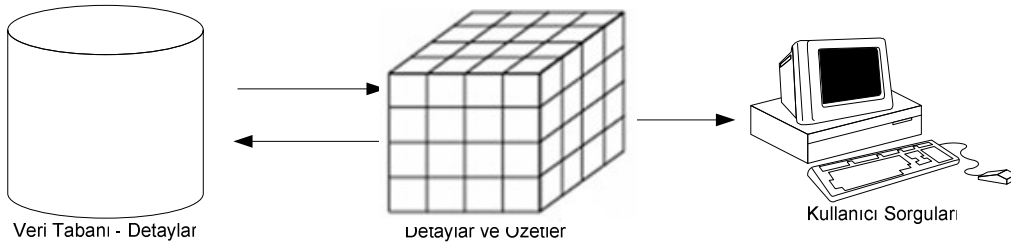
Şekil 2.3 ROLAP Mimarisi

**MOLAP** (Multidimensional OLAP), çok boyutlu veri küplerin veritabanı haricinde dış bir kaynak üzerinde hazırlanması sağlar. Çok boyutlu verilerin dizi gibi özel veri yapıları üzerinde tutulmasını sağlar ve ÇİAİ işlemleri bu özel veri yapılarına erişilerek gerçekleştirilir. Tanımlanmış hesaplamalar için veriler özetlenir ve işletilir. MOLAP depolama mekanizması, verinin çekilmesinde oldukça hızlıdır. İlişkisel veritabanlarında sorgulara cevap vermek için ilgili sorguya ait sonuç kümesi oluşturulurken, MOLAP hesaplanmış verilerin sistemdeki konumunu sakladığı için ilgili adrese giderek verinin çekilmesini sağlar. MOLAP yönteminin kötü tarafı, küpün veri ambarı ile bağlantısının koparılmış olmasıdır. Veri ambarı bir ilişkisel veritabanı kapsamında yer alırken, küpler özel veri yapıları olarak sistemde saklanır. Bu durum küpün güncellenmesi işlemini oldukça karmaşıklaştırır.



Şekil 2.4 MOLAP Mimarisi

**HOLAP** (Hybrid OLAP) Her iki yöntemi de içinde barındırır. Veriler VTYS üzerinde, verilere ait kümeleme sonuçları özel veri yapıları üzerinde saklanır. Bazı durumlarda %80 yüksek seviyeli özetlenmiş veriye gereksinim duyan sorgular gerekirken, %20 oranında düşük seviyede kaynak verideki satırlara karşılık gelen sorgular olabilir. Bu tür durumlar için HOLAP yaklaşımı uygundur. HOLAP hem MOLAP hem de ROLAP yaklaşımlarının avantajlarını bir arada bulundurur. Yüksek seviyeli özetlenmiş veriler MOLAP küpleri ile saklanırken, düşük seviyeli özet verileri ve tablolardaki kayıtlara ilişkisel veritabanından erişim sağlanır. ÇİAİ motoru çalıştırılacak sorguya göre farklı saklama mekanizmalarında bulunan verileri seçebilme yeteneğine sahiptir. HOLAP kapsamında veritabanında tutulacak tabloların karmaşıklığı olabildiğince azdır. Dolayısıyla, kurulacak indeksleme mekanizmaları ile veri olabildiğince eniyelenir. Özetlenmiş veriler üzerinden yapılan sorgular MOLAP mekanizması ile ele alındığı için oldukça hızlıdır. Oracle Express Server hem ilişkisel hem de çok boyutlu verileri yönetebilen HOLAP yaklaşımına göre çalışan bir sunucu yazılımıdır.



Şekil 2.5 HOLAP Mimarisi

## 2.7. Veri Ambarlarından Çok Boyutlu Verinin Elde Edilmesi

İlişkisel veritabanları normal veri yapıları üzerinde yapılan işlemlerde oldukça esnekler. Günlük işletimsel işlemlerin yürütüldüğü ÇİHİ sistemlerinde küçük hareketler ve birincil anahtar ve indeksler üzerinden yapılan sorgular desteklenir. ÇİHİ sistemlerinde anlık bilgiler vardır fakat veri ambarları tarihsel ve özet verileri kullanır. ÇİAİ teknolojisi karar destek kapsamında yöneticilere bilgi sağlamak için veri ambarlarında kullanılan bilgi yapılarıdır.

Veri ambarları veri analizinde kullanılan şemalarını ÇİAİ araçları gibi çeşitli araçlarla oluştururlar.

ÇİAİ araçları, verinin çok boyutlu kavramsal görünümünü desteklemektedir. Çok boyutlu modelde, veriler satır ve sütunlar yerine olay (fact) ve boyutlarda (dimension) saklanmaktadır. Verinin çok boyutta modellenmesi ve incelenmesi “Veri Küpleri” ile sağlanır.

Veri küplerini oluşturan temel öğeler aşağıdakilerdir:

- Olaylar
- Boyutlar
- Ölçüm Değerleri

### **2.7.1. Olaylar**

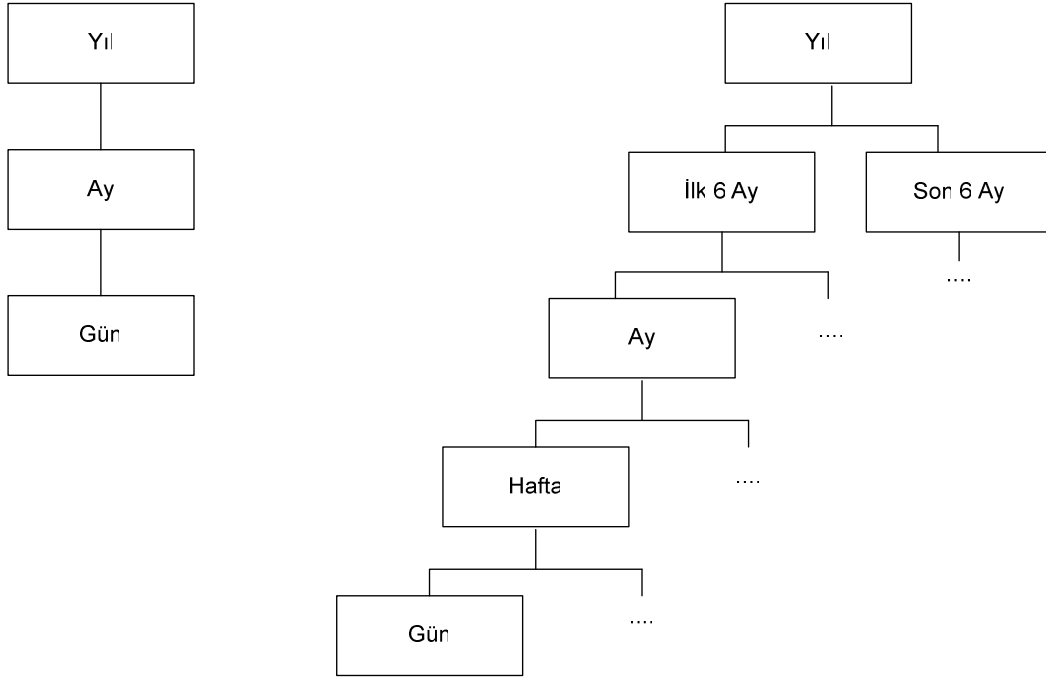
Olay, ölçülebilir nicel değerlere sahip konu bütünlüğü içerisindeki bir olgudur. Veri ambarı kapsamında olaylar, olay tablo ile ifade edilir. Olay tablo, sayılabilir nicel verileri içeren çok boyutlu veriyi tutar. Örneğin, ‘Muayene’ tablosu olay tablodur ve belirli bir tarihte, belirli bir uzmanlık alanı kapsamında yapılmış olan muayene işlemi sayılarını saklar.

### **2.7.2. Boyutlar**

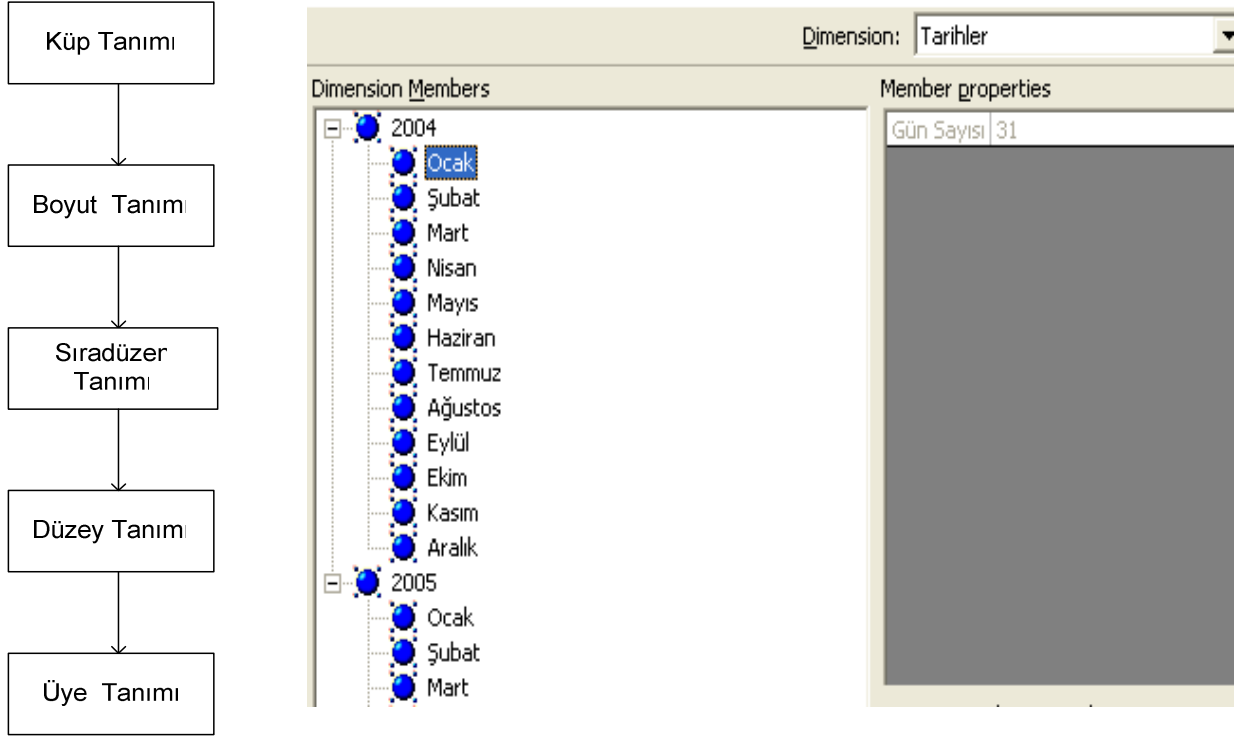
Boyut, genel olarak bir çizelge üzerindeki bir ekseni ifade eder. Boyut tablo, olay tabloya tanımlayıcı nitelikler ile ilişkili tablodur. Boyut veri üzerinde tekil bir bakış açısını simgeler. Tekil ölçüm değeri ya da küpün bir yüzüne karşılık gelir. Boyutlar sayısal olmayan fakat bilgi verici varlıklar olarak tanımlanabilir. Örneğin, ‘Muayene’ olayı kapsamında ‘Tarih, Bölümler, Uzmanlık ve Hasta Durumu’ boyutları oluşturulur. Her bir boyut bilgisi veri ambarında boyut tablosu olarak saklanır.

Boyut üyeleri, boyut içerisinde yer alan her bir veri ögesini temsil eder. ‘Tarih’ boyutu için yıl bilgisi olarak [2004] değeri bir boyut elemanı olduğu gibi [2004].[Ocak] değeri 2004 yılının Ocak ayını temsil eden başka bir boyut elemanıdır. Benzer şekilde ‘Hasta Durumu’ boyutu için [Ayaktan] ve [Yatan] bilgi içerikleri ilgili boyutun boyut elemanlarını ifade eder.

Boyutlar bir ya da daha fazla sıradüzensel yapıya sahip olabilirler. Örneğin, 'Tarih' boyutu 'yıl – ay – gün' sıradüzenine sahip olabileceği gibi aynı zamanda 'yıl – çeyrek – ay – hafta – gün' sıradüzenine de sahip olabilir. Bu örnek kapsamında tanımlanan yeni sıradüzen özellikle mali takvim çerçevesinde haftalara göre bölünme gereksiniminin olduğu durumlarda kullanılır. Bu durum Şekil 2.6 ile gösterilmiştir.



Şekil 2.6 Boyut sıradüzen gösterimi



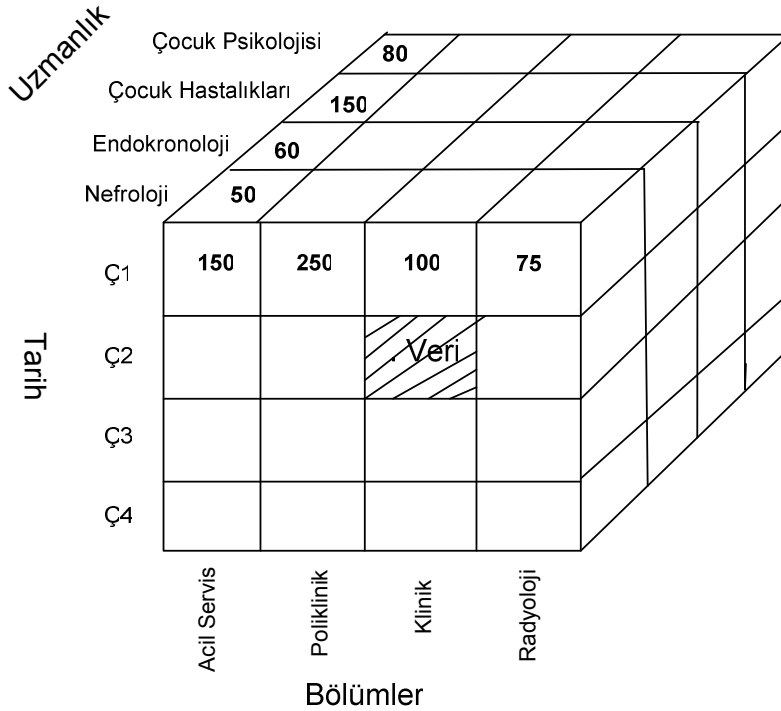
Şekil 2.7 Küp – Üye Tanımı Akışı

Şekil 2.7 üzerinde belirtildiği gibi küpler boyutlardan, boyutlar bir ya da daha fazla sıradüzensel yapıdan, her bir sıradüzen düzeylerden ve her düzey de üyelere oluşur. Bir boyut bilgisi, düzey seviyesinde kalabilir, yani her düzeyde üye bulunması gerekmemektedir. Bu durumu Şekil 2.7’de belirtilen “Tarihler” boyutu ile örnekleyelim. “Tarihler” boyutunda “Yıl – Ay” şeklinde bir sıradüzen yapısı bulunmaktadır. Tanımlanmış olan bu sıradüzene göre her bir boyut üyesi uygun düzeylerde yer almaktadır. 2004 yılının altında Ocak , Şubat, ..., Aralık şeklinde aylar düzeylendirilmiştir. Bu düzeylendirme diğer yıllar için de tanımlanmış sıradüzensel yapıya göre devam etmektedir. Her bir ay bilgisi için ilgili ayın yıl içerisinde kaç güne sahip olduğu bilgisi “Gün Sayısı” şeklinde adlandırılmış üye tanımı ile yapılmaktadır. Verilen şekilde 2004 yılı Ocak ayı için 31 değeri üye tanımı ile tutulmaktadır.

### 2.7.3. Ölçüm Değerleri

Ölçüm değeri, bir olayın sayısal niteliğidir. Bu nitelik boyut değerlerine göre anlam kazanır. Her bir boyutun kesişim noktasında yer alan hücre, ölçüm verisidir.

Şekil 2.8 üzerinde “Muayene Küpü” gösterilmektedir. Muayene sayılarının yer aldığı “Muayene” tablosu olay tablodur. Bu tabloya tanımlayıcı niteliklerle ilişkili olan “Tarih, Bölümler ve Uzmanlık” tabloları boyutları oluşturur. Boyutların kesişiminde yer alan hücre üzerindeki veri muayene sayısını ifade eden ölçüm değeridir. Şekil üzerinde belirtilen örneğe göre, Poliklinik bölümünde yılın ilk çeyreğinde 250 , Endokronoloji uzmanlık dalında Acil Servis bölümünde 60 adet muayene işleminin gerçekleştiğini görebiliyoruz.



Şekil 2.8 Muayene Küpü Gösterimi

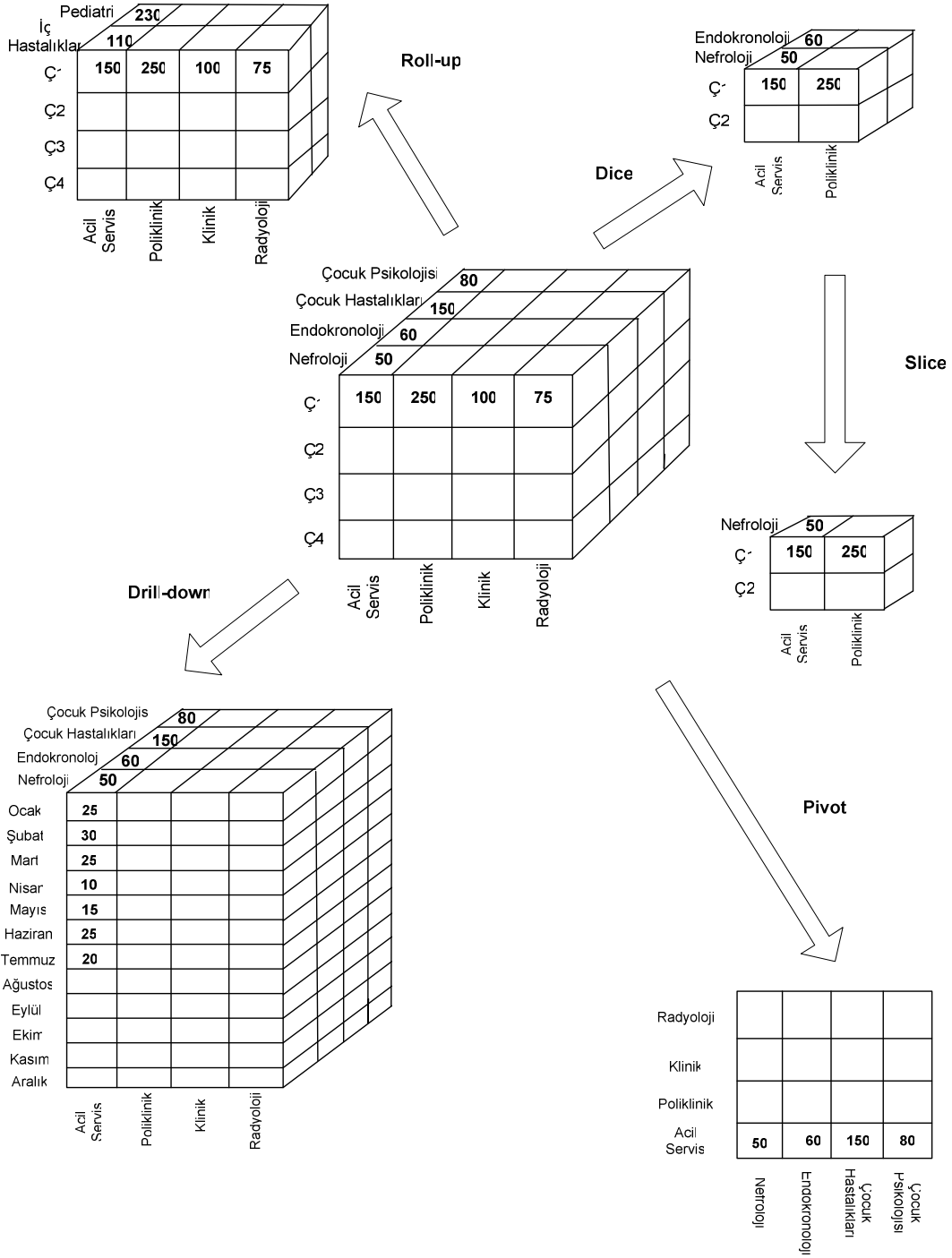
Olay tablo, genellikle büyük yoğunlukta veri taşır. Olay ve boyut tablolarından oluşmuş bir küpe ait verilerin hepsine eş zamanlı olarak erişmek ve bu verileri doğru olarak algılayabilmek mümkün değildir. Bu nedenle, küp üzerinde ikişer boyut seçilerek verinin anlaşılabilirliği ve kullanıcıya erişim hızı artırılır. Veriye erişimlerde iyi bir performansı yakalayabilmek için tek bir olay tabloya bağlanan



boyutların sayısını yönetilebilir bir seviyede tutmak gerekir. Boyutların sayısı arttıkça, olay tablo üzerindeki satırların sayısı da artacaktır.

ÇİAİ istemcileri tarafından gerçekleştirilen genel işlemler; (Chen Z., 2001)

- “**Roll-up**”, boyutlar üzerinde gruplama seviyesini artırarak daha özet veriye ulaşmayı sağlar. Verinin özetlenmiş haline ulaşılır. Şekil 2.9 üzerinde uzmanlık alanları olarak ‘Çocuk Psikolojisi, Çocuk Hastalıkları, Endokronoloji ve Nefroloji’ yer almaktadır. ‘Roll-up’ işlemi sonrasında uzmanlık alanları gruplanarak ‘Çocuk Psikolojisi ve Çocuk Hastalıkları’ ‘Pediatri’ olarak, ‘Endokronoloji ve Nefroloji’ ‘İç Hastalıklar’ olarak gruplanmış ve bu gruplar altında daha özet verinin sunulması sağlanmıştır.
- “**Drill-down**”, daha yüksek ayrıntı düzeylerine inilir , ‘roll-up’ işleminin tersidir, gruplama seviyesi azaltılarak çok daha detaylı bilgilere erişilir. Şekil 2.9 üzerinde tarih boyutu yıl içerisindeki çeyreklere göre gruplanmıştır. ‘Drill-down’ işlemi sonrasında her bir çeyrek kapsadığı ay bilgilerine göre genişletilerek verinin daha detaylı gösterimi sağlanmıştır.
- “**Slice ve dice**”, seçme ve izdüşüm almadır, verinin boyutu azaltılıp, belirli bir boyutun alt kümesi alınarak izdüşüme erişilir. Şekil 2.9 üzerinde ‘dice’ işlemi sonrasında uzmanlık alanı boyutundan ‘İç Hastalıklar’ bölümü, tarih boyutundan ilk iki çeyrek ve bölümler boyutundan ‘Acil Servis ve Poliklinik’ boyut üyeleri seçilerek bakış açısı belirli bir kümeye göre daraltılmıştır. Küp üzerinde döndürme işlemi yapılarak tüm açılardan veriye erişim sağlanır. ‘Slice’ işlemi sonrasında küp dilimlere ayrılarak sadece ‘Nefroloji’ uzmanlık alanına ait dilimine odaklanma sağlanmıştır.
- “**Pivot**”, çok boyutlu görüntünün yeniden yönlendirilmesi işlemidir. Çok boyutlu verinin (x,y) koordinatlarıyla ifade edilecek alt parçalara bölünmesini sağlar. Boyutlardan sadece ikisi kullanılır ve boyutların kesişim noktası ölçüm değerini verir. Çok boyutlu verinin iki boyutlu olarak incelenmesini sağlar.



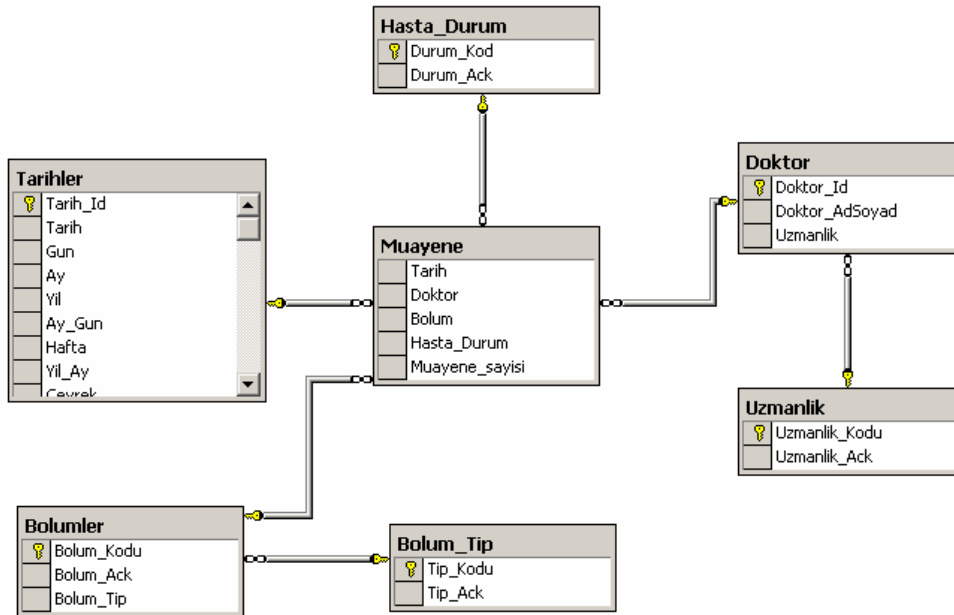
Şekil 2.9 ÇİAl İşlemleri Çizimsel Gösterimi

## 2.8. Veritabanı Tasarım Metodolojileri

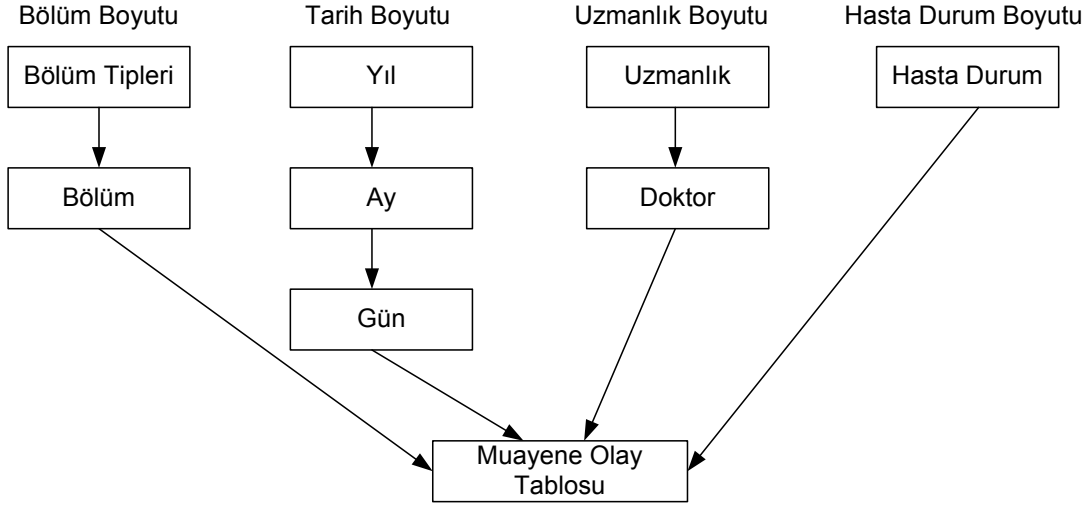
ÇİHi sistemlerinde varlık-bağıntı çizenekleri ile veritabanı tasarımı yapılmaktadır. Fakat çok yoğun sorgulama ve veri yükleme içeren karar destek sistemleri için varlık-bağıntı çizenekleri çok uygun değildir.

Veritabanı tasarımında, ÇİAi sorgu motorları ve ön uç araçları etkileyen en popüler kavramsal model çok boyutlu görüntülerdir. Çok boyutlu modelde istenilen sayıda ölçüm değerini içeren konu başlıkları vardır. Ameliyat, muayene, tetkikler bu tür konu başlıklarına örnek olarak verilebilir. Her konu başlığı bir boyut kümesine bağlıdır. Örneğin uzmanlık dalı, bölüm ve tarih muayene konusunun boyutları olabilir. Her boyut, nitelik kümesiyle tanımlanır. Boyuta ait nitelikler, sıradüzensel bir yapıda tutulur.

Çok boyutlu modelde, varlık-bağıntı çizeneğinde yer alan varlıklar boyutlarla eşleştirilir. Birden çok boyuta sahip olan ve ölçülebilir nicel değerleri barındıran varlık, olay tablo olarak nitelendirilir. Şekil 2.10 ile 'Muayene' küpüne karşılık gelen varlık-bağıntı çizeneği ve Şekil 2.11 ile bu çizeneğin çok boyutlu model kapsamındaki ifadesi yer almaktadır.



Şekil 2.10 Muayene İşlemini Gösteren Varlık-Bağıntı Çizeneği



Şekil 2.11 Varlık-Bağıntı Çizeneğinin Çok Boyutlu Modele Dönüştürümü

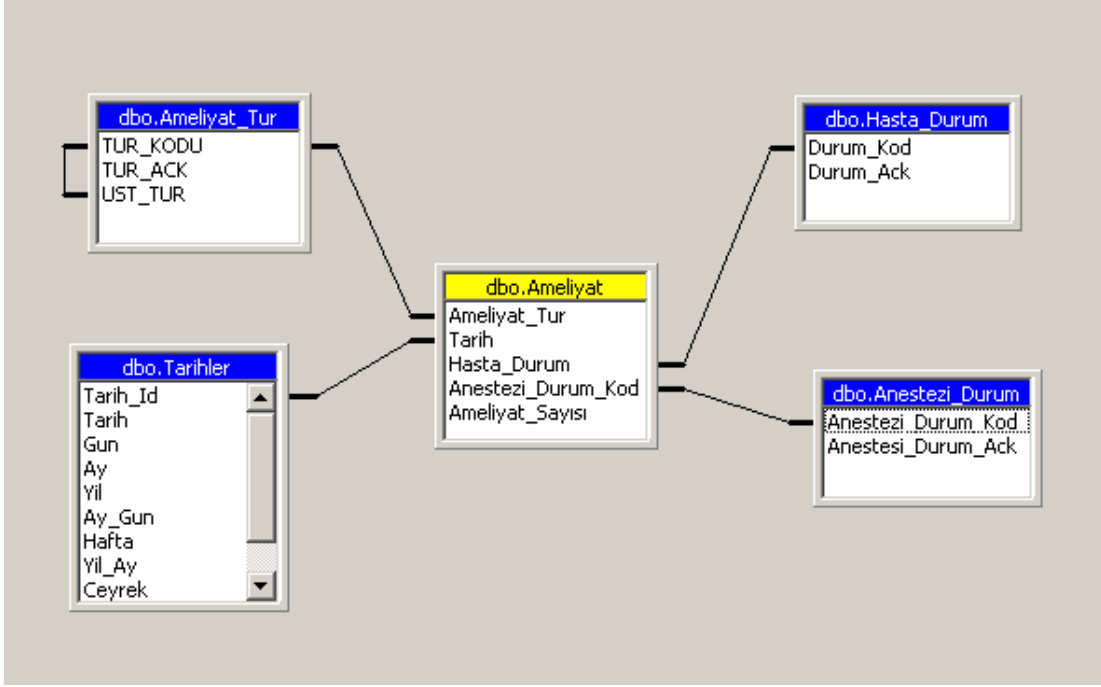
Veri ambarı tasarımında

- Yıldız Şema Modeli
- Kar Yağışı Şema Modeli

olmak üzere iki yöntem mevcuttur.

### 2.8.1. Yıldız Şema

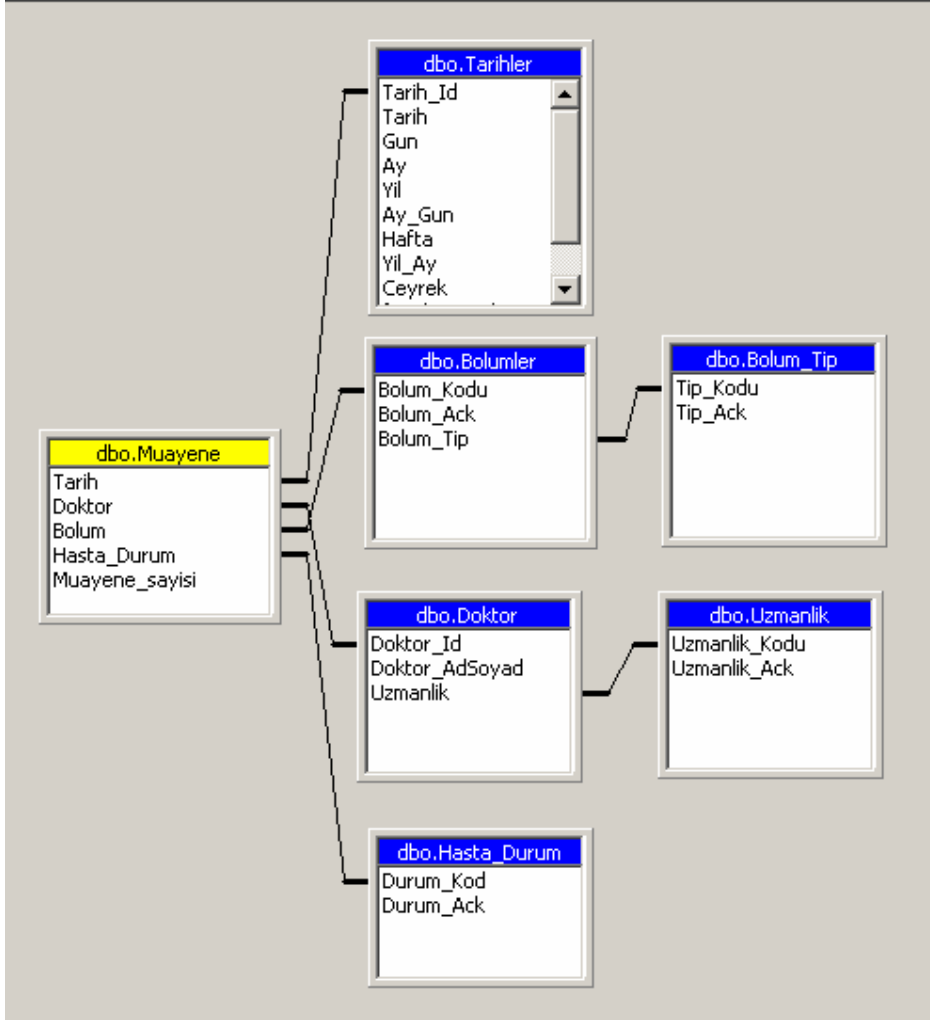
Birçok veri ambarı, 'yıldız şema' yöntemini kullanarak çok boyutlu verileri modellemektedir. Veri tabanında bir tane olay tablo ve her bir boyut için bir tane boyut tablo yer alır. Boyut tabloları olay tablosu etrafında yıldız şeklinde sıralandığı için 'yıldız şema' olarak adlandırılmıştır. Her boyut tablosunda, boyuta ait nitelikleri tutan kolonlar yer alır. Olay tablosundaki her çoklu her bir boyut tablosunu referans veren göstergeler içerir. Şekil 2.12 üzerinde 'Ameliyat' küpüne ait yıldız şema gösterimi yer almaktadır. Nicel sayılabilir nitelik olarak ameliyat sayılarını tutan 'Ameliyat' olay tablosu etrafında yıldız şeklinde dağılmış 'Ameliyat Türü, Hasta Durumu, Tarih ve Anestezi Durumu' boyut tabloları yer almaktadır. Boyut tablolarında bulunan birincil anahtar konumundaki sütunlar olay tablosunda bulunan ilgili sütunlarla ilişkilendirilmiştir. Ayrıca her bir boyut için tek bir boyut tablosu bulunmaktadır.



Şekil 2.12 Ameliyat Küpüne Ait Yıldız Şema Gösterimi

### 2.8.2. Kar Yağışı Şeması

'Kar yağışı' modelinde, boyut tabloları birbiriyle yabancı anahtar üzerinden bağlı yeni boyut tabloları ile ilişkili olurlar. Bir boyut tanımlamak için birden çok boyut tablosu kullanılabilir. Bu tablolar arasında birçoğa-bir türü bir ilişki mevcuttur. Bu ilişki kapsamında sağlanan sıradüzene sahip boyut yapısı tanımlanır. Şekil 2.13 üzerinde 'Muayene' küpünü gösteren kar yağışı şeması bulunmaktadır. Bu şemada yer alan, 'Bölüm ve Uzmanlık' boyutları iki tablo kullanılarak tanımlanmıştır. 'Bölüm' boyutunda, bölüm tiplerine göre bölümler sıradüzensel olarak tutulmaktadır. Örneğin , 'Poliklinik, Klinik, Laboratuar' bölüm tiplerini oluştururken, 'Mikrobiyoloji, Patoloji, Genetik Laboratuvarı' 'Laboratuar' bölüm tipine sahip bölümleri belirtmektedir. Dolayısıyla muayene küpü üzerinden yapılacak sorgularda her bir bölüm tipine ve ilgili tipe sahip bölümlerde yapılmış olan muayene sayıları sıradüzensel olarak elde edilebilecektir.



Şekil 2.13 Muayene Kuponüne Ait Kar Yağışı Gösterimi

## 2.9. Çok Boyutlu İfadeler – ÇBİ (Multidimensional Expressions-MDX)

Tez kapsamında kullandığım uygulama ortamı Microsoft SQL Server veri tabanı ve MSSQL Server OLAP Service ÇİAİ servisedir. Bu kapsamda Microsoft Analysis Manager ÇİAİ aracının imkân ve kabiliyetlerine göre çalışmamı tamamladım. Microsoft ürün ailesinin ÇİAİ kapsamında hazırlanan sorguları tanımlayan ifadeler için kullandığı yöntem 'Çok Boyutlu İfadeler' dir. Bu ifadeler, yapısal sorgulama diline yakın olsa da bir çok yönden farklılıklar göstermektedir. Özel veri yapıları altında saklanan küp verilerine erişerek istenilen kriterlere göre sonuçlara ulaşılması sağlanır. ÇBİ, Applix, Microsoft, Microstrategy, SAS, SAP ve Whitelight sunucuları tarafından desteklenmektedir.

### 2.9.1. ÇBİ Sözdizim Yapısı

ÇBİ kullanılarak çok boyutlu küpler, boyutlar, seviyeler yaratılabilir, mevcut küp yapıları güncellenebilir, küplere veriler eklenebilir ve mevcut veriler sorgulanabilir. Tez kapsamında sorgulama amaçlı raporlama aracı geliştirildiği için sadece sorgulama amaçlı kullanılan sözdizim yapıları üzerinde durulacaktır.

ÇBİ ile çok boyutlu küplerde saklanan verilerin iki boyut üzerinde sorgulanmasını sağlayacak ifadeler üretilir. (Spofford G., 2001)

#### 2.9.1.1 Genel Sözdizim Yapısı

Sorgulama amaçlı en genel sözdizim yapısı;

#### **SELECT**

{ ..... } **on columns,**

{ ..... } **on rows**

#### **FROM** Küp\_İsmi

'Select, From, on columns ve on rows' terimleri anahtar kelimelerdir. 'on columns' terimleri öncesinde x eksenini üzerinde, 'on rows' terimleri öncesinde de y eksenini üzerinde yer almasını istediğimiz boyut değişkenleri verilir. Köşeli parantezler içerisine sırasıyla boyutun adı, sıradüzeni, seviyesi ve özellikleri yazılır. Eğer boyut tek bir sıradüzenden oluşuyorsa boyut adından sonra doğrudan seviyeye geçilir. İlgili boyuta ait tüm seviyeler görüntülenmek isteniyorsa sadece boyut adı verilir.

Örneğin, 'Tarih' boyutunu ele alalım. 'yıl-ay-gün' şeklinde seviyelerimiz bulunsun.

[Tarih].members ifadesi, tarih boyutundaki tüm öğelerin elde edilmesini sağlar.

[Tarih].[2004].children ifadesi tarih boyutunun yıl seviyesine göre 2004 yılına ait ay ve gün seviyelerine ulaşımı sağlayacaktır.

[Tarih].[2004].[Ocak] ifadesi, tarih boyutunun 2004 yılının Ocak ayına ait bilgilerine erişim sağlayacaktır.

“Select { [Tarih].[2004].children } on columns , { [Hasta\_Durum].members } on rows From Ameliyat “ ifadesi ile ameliyat küpü kapsamında, x ekseninde tarih boyutunun yıl seviyesine göre 2004 yılına ait değerleri, y ekseninde hasta durum boyutunun tüm elemanları yer alacaktır ve x ve y eksenlerinin kesişim noktasını veren hücreler üzerinde iki boyut değerine karşılık gelen ölçüm değerleri görüntülenecektir.

## **SELECT**

{ ..... } **on columns,**

{ ..... } **on rows**

**FROM** Küp\_İsmi

**WHERE** ( ..... )

ÇBİ kapsamında ‘where’ anahtar kelimesinin kullanımı ile elde edilecek sorgunun çeşitli kısıtlarla sınırlandırılması mümkündür. Örneğin;

“Select { [Tarih].[2004].children } on columns , { [Hasta\_Durum].members } on rows From Ameliyat Where ( [Anestezi].[yapıldı] )” ifadesi ile, ameliyat küpü kapsamında sadece anestezi uygulanmış ameliyatların, x ekseninde 2004 yılına ait tarih boyutu ve y ekseninde hasta durum boyutunun görüntülenmesini sağlayacaktır.

### **2.9.1.2 Küme ve Çoklu İşlemleri**

x ve y ekseninde yer alacak boyut içerikleri kümeler ve çoklular şeklinde detaylandırılabilir.

**Virgül (,)**, kullanılarak birden çok boyut elemanının yan yana gelecek şekilde birleştirilerek sorgulanması sağlanır. ‘and- ve’ işlecine karşılık gelir. Örneğin,

“Select { [Tarih].[2004].[Ocak], [Tarih].[2004].[Mart], [Tarih].[2004].[Haziran] } on columns , { [Hasta\_Durum].members } on rows From Ameliyat Where ([Anestezi].[yapıldı] )” ifadesi ile, ameliyat küpü üzerinde tarih boyutunun 2004 yılı



içerisinde sadece Ocak, Mart ve Haziran aylarına ait değerlerinin x ekseninde yer alması sağlanır.

**İki nokta (:)** işleci ile verilen boyut elemanları arasında kalan tüm elemanların sorgulanması sağlanır. 'between-arasında' işlecine karşılık gelir. Örneğin,

“Select { [Tarih].[2004].[Ocak] : [Tarih].[2004].[Haziran] } on columns , { [Hasta\_Durum].members } on rows From Ameliyat Where ( [Anestezi].[yapıldı])” ifadesi ile, ameliyat küpü üzerinde tarih boyutunun 2004 yılı içerisinde sadece Ocak ve Haziran ayları arasındaki tüm aylara ait değerlerinin x ekseninde yer alması sağlanır.

**'members'** ifadesi ile bu ifadenin solunda yer alan boyut, sıradüzen ya da seviyenin tüm elemanlarına ait çokluların elde edilmesi sağlanır.

**'children'** ifadesi ile bu ifadenin solunda yer alan boyut, sıradüzen ya da seviyenin çocuklarına ait çokluların elde edilmesi sağlanır.

**'Union'** işleci ile belirtilen kümelerin birleşimine karşılık gelen çokluların sorgulanması sağlanır. Örneğin,

“Select { Union ([[Tarih].[2004].[Ocak] : [Tarih].[2004].[Haziran] ] , [[Tarih].[2004].[Ocak] , [Tarih].[2004].[Ağustos]]) } on columns , {[Hasta\_Durum].members} on rows From Ameliyat Where ( [Anestezi].[yapıldı])” ifadesi ile 2004 yılında Ocak ve Haziran ayları arası ve Ağustos ayına ait boyut değerlerinin elde edilmesi sağlanır.

Aynı değer aralığını ifade eden bölümler tekrar edilmez. Yukarıdaki ifade de 'Ocak' ayı her iki küme içerisinde de olmasına rağmen sorgu sonucunda bir kez görüntülenir. 'Union' işleci en fazla iki küme üzerinde birleştirme işlemine izin verir. İki künden çok kümenin birleştirilmesi için iç içe 'Union' ifadeleri kullanılmalıdır.

**'Intersect'** işleci ile, belirtilen kümelerin kesişimine karşılık gelen çokluların sorgulanması sağlanır. Örneğin,

“Select { Intersect ( [[Tarih].[2004].[Ocak] : [Tarih].[2004].[Haziran] ] , [[Tarih].[2004].[Ocak] , [Tarih].[2004].[Ağustos]]) } on columns ,

{[Hasta\_Durum].members} on rows From Ameliyat Where ( [Anestezi].[yapıldı])” ifadesi ile 2004 yılında Ocak ayına ait boyut değerlerinin elde edilmesi sağlanır.

‘Intersect’ işleci en fazla iki küme üzerinde birleştirme işlemine izin verir. İki den çok kümenin birleştirilmesi için iç içe ‘Intersect’ ifadeleri kullanılmalıdır.

### 2.9.1.3 Kartezyen Çarpım (CrossJoin)

Bazı durumlarda farklı iki küme üzerinde kartezyen çarpım sonucuna ihtiyaç duyulur. Bu sayede aynı anda sadece iki boyutunu sorgulayabildiğimiz çok boyutlu görüntülerde erişebileceğimiz boyut sayısı artmaktadır. Bunun için sorgulanacak kümeler ‘Crossjoin’ işleci ile tek bir küme olarak ifade edilir. Örneğin,

“select { Crossjoin ( [Tarih].[2004].[Ocak] : [Tarih].[2004].[Aralık] , [Hasta\_Durum].members )} on columns, { [Anestezi].members } on rows from Ameliyat” ifadesi ile x ekseninde her ay için hasta durum bilgisi tekrarlanmaktadır. Bu durumu ifade eden grid yapısı Şekil 2.14 ile gösterilmiştir. ‘CrossJoin’ işleci en fazla iki küme üzerinde kartezyen çarpım işlemine izin verir. İki den çok kümenin kullanılması durumunda iç içe ‘CrossJoin’ ifadeleri kullanılmalıdır.

|           | Ağustos |         |       | Eylül |         |       | Ekim |         |       | Kasım |         |
|-----------|---------|---------|-------|-------|---------|-------|------|---------|-------|-------|---------|
|           | Tümü    | Ayaktan | Yatan | Tümü  | Ayaktan | Yatan | Tümü | Ayaktan | Yatan | Tümü  | Ayaktan |
| Tümü      |         |         |       | 14,00 | 6,00    | 8,00  | 6,00 | 4,00    | 2,00  | 8,00  | 8,00    |
| Yapıldı   |         |         |       | 14,00 | 6,00    | 8,00  | 6,00 | 4,00    | 2,00  | 8,00  | 8,00    |
| Yapılmadı |         |         |       |       |         |       |      |         |       |       |         |

Şekil 2.14 Kartezyen Çarpım Grid Örneği

### 2.9.1.4 Filtreleme (Filter)

Belirtilen kriterlere göre sonuç kümesinin daraltılması için kullanılan işleçtir. ‘Filter’ işleci parametre olarak bir küme ve bir mantıksal ifadeye sahiptir ve mantıksal ifadenin ‘Doğru’ olduğu durumları kapsayan kümeyi geri döndürür. Örneğin,

“select { Filter ( {Crossjoin( [Tarih].[2004].[Ocak] : [Tarih].[2004].[Aralık] , [Hasta\_Durum].members )}, [Measures].[Ameliyat Sayısı] > 4)} on columns, {[Anestezi].members } on rows from Ameliyat “ ifadesi ile ameliyat sayısı 4 ten

büyük olan kayıtların sorgulanması sağlanır. Bu ifadenin grid üzerindeki görüntüsü Şekil 2.15 ile gösterilmiştir.

|           | Eylül |         |       | Ekim | Kasım |         | Aralık |
|-----------|-------|---------|-------|------|-------|---------|--------|
|           | Tümü  | Ayaktan | Yatan | Tümü | Tümü  | Ayaktan | Tümü   |
| Tümü      | 14,00 | 6,00    | 8,00  | 6,00 | 8,00  | 8,00    | 6,00   |
| Yapıldı   | 14,00 | 6,00    | 8,00  | 6,00 | 8,00  | 8,00    | 6,00   |
| Yapılmadı |       |         |       |      |       |         |        |

Şekil 2.15 Filtreleme Grid Örneği

### 2.9.1.5 Sıralama (Order)

'Order' işleci, orijinal kümeye ait çokluların belirtilen şekilde sıralanmasını sağlar. Sıralanacak olan küme, sıralanacak kümeye ait kriterler ve sıralamanın azalan ya da artan sırada olduğunu belirten anahtar kelime kullanılır. Örneğin,

“ Order ( [Tarih].members , [Measures].[Ameliyat Sayısı], BDESC) “ ifadesi ile tarih boyutuna ait öğeler, ilgili tarihte gerçekleşmiş ameliyat sayılarına göre azalan içerikte sıralanır.

### 2.9.1.6 Hesaplanan Öğeler (Calculated Members)

ÇBİ ile sadece boyut, hiyerarşi ve seviye üzerinden sorgulama yapılma zorunluluğu yoktur. Aynı zamanda, hesaplanan öğeler tanımlanarak formülize edilmiş çeşitli satır ve sütun elemanları oluşturulabilir. Bunun için ölçüm değerleri üzerinden çeşitli matematiksel ifadeler geliştirilebilir ya da ÇBİ kapsamında tanımlanmış olan çeşitli işlevler (PeriodsToDate, ParallelPeriod, YTD, Sum, ...vb) kullanılarak istenilen değerlere erişim sağlanabilir. Hesaplanan üye tanımlanması için “ with member ” anahtar sözcüğü ile öncelikle bu üneyi ifade edecek biricik bir isim ve bu ünenin ifade edeceği veri kümesini kapsayacak ifadenin belirtilmesi gerekir. Daha sonra belirtilen isim ile x ya da y ekseninde görüntülenecek şekilde ÇBİ içerisine yerleştirilir. Örneğin,

“with member [Anestezi].[Anestezi Fark] As '[Anestezi].[Yapıldı]-[Anestezi].[Yapılmadı]'

select { Crossjoin( [Tarih].[2004].[Ocak] : [Tarih].[2004].[Aralık] ,  
 [Hasta\_Durum].members )) on columns, {[Anestezi].members ,  
 [Anestezi].[Anestezi Fark] } on rows from Ameliyat “ ifadesi ile anestezi uygulanmış  
 ameliyatlara ile uygulanmamış ameliyat sayıları arasındaki fark ‘Anestezi Fark’  
 olarak adlandırılmış ve formüle edilmiş ifadesi ‘[Anestezi].[Yapıldı]-  
 [Anestezi].[Yapılmadı]’ şeklinde tanımlanmıştır. Tanımlanan bu ifade y eksenini  
 kapsamında seçilerek sorgu sonucunda görüntülenmesi sağlanmıştır. Bu ifadenin  
 grid üzerindeki görüntüsü Şekil 2.16 ile gösterilmiştir.

|               | Eylül   |       | Ekim |         |       | Kasım |         |       | Aralık |         |
|---------------|---------|-------|------|---------|-------|-------|---------|-------|--------|---------|
|               | Ayaktan | Yatan | Tümü | Ayaktan | Yatan | Tümü  | Ayaktan | Yatan | Tümü   | Ayaktan |
| Tümü          | 6,00    | 8,00  | 6,00 | 4,00    | 2,00  | 8,00  | 8,00    |       | 6,00   | 4,00    |
| Yapıldı       | 6,00    | 8,00  | 6,00 | 4,00    | 2,00  | 4,00  | 4,00    |       | 6,00   | 4,00    |
| Yapılmadı     |         |       |      |         |       | 4,00  | 4,00    |       |        |         |
| Anestezi Fark | 6,00    | 8,00  | 6,00 | 4,00    | 2,00  | 0,00  | 0,00    |       | 6,00   | 4,00    |

Şekil 2.16 Hesaplanan Öğe Grid Örneği

“with

member [Measures].[Ay Fark] as '([Measures].[Eylül 2005]-[Measures].[Eylül 2004])'

member [Measures].[Eylül 2004] as '([Measures].[Eylül 2005], ParallelPeriod([Yil], 1))'

member [Measures].[Eylül 2005] as 'Sum(PeriodsToDate([ay],[tarihler].currentmember),[Measures].[Ameliyat Sayısı])'

member [Measures].[Yıl Fark] as '([Measures].[2005]-[Measures].[2004])'

member [Measures].[2004] as '([Measures].[2005], ParallelPeriod([Yil], 1, [tarihler].parent))'

member [Measures].[2005] as 'Sum(PeriodsToDate([yil],[tarihler].currentmember),[Measures].[Ameliyat Sayısı])'

select { [Measures].[Eylül 2004] , [Eylül 2005] , [Measures].[Ay Fark] ,  
 [Measures].[2004] , [Measures].[2005] , [Measures].[Yıl fark] } on columns,  
 {[Hasta\_Durum].members} on rows from Ameliyat where ([tarihler].[2005].[Eylül])“  
 ifadesi ile altı tane hesaplanan öğe tanımlanmıştır. Bu öğeler ÇBİ'nin “where”  
 bölümünde belirtilen tarih kısıtına göre çeşitli aralıklardaki ameliyat sayılarının  
 toplamlarını üretmektedir. ‘[Measures].[2005]’ öğesi, 2005 yılındaki tüm ameliyat  
 sayısını üretmektedir. ‘[Measures].[2004]’ öğesi, 2004 yılında gerçekleşmiş

ameliyatların sayılarını üretmektedir. '[Measures].[Yıl Fark]' ögesi, 2004 ve 2005 yıllarında gerçekleşmiş ameliyat sayılarının farkını vermektedir. '[Measures].[Eylül 2005]' ögesi 2005 yılı Eylül ayı, '[Measures].[Eylül 2004]' ögesi 2004 yılı Eylül ayında gerçekleşmiş ameliyat sayılarını üretmektedir. Bu ifadenin grid üzerindeki gösterimi Şekil 2.17 ile gösterilmiştir.

|         | Eylül 2004 | Eylül 2005 | Ay Fark | 2004  | 2005  | Yıl Fark |
|---------|------------|------------|---------|-------|-------|----------|
| Tümü    | 14,00      | 42,00      | 28,00   | 34,00 | 80,00 | 46,00    |
| Ayaktan | 6,00       |            | -6,00   | 22,00 |       | -22,00   |
| Yatan   | 8,00       | 42,00      | 34,00   | 12,00 | 80,00 | 68,00    |

Şekil 2.17 Hesaplanan Öge Grid Örneği

### 3. VERİ MADENCİLİĞİ

Veri madenciliği, büyük yoğunluktaki tarihsel veri üzerinde analiz ve kümeleme sonucunda, anlamlı örüntülerin ve ilişkilerin keşfedildiği işlemdir (Seidman, 2001). Veri madenciliğini, organizasyonların karar aşamaları için yeni bilgiler üreten ya da gelecekle ilgili tahminler ve planlar yapmamızı sağlayan bir dizi teknikler ve anlayışlar bütünü olarak da tanımlayabiliriz.

Madencilik teriminin kullanılma sebebi, büyük bir veri yığını arasından uygun olanı arama ve seçme işleminin maden arama işine benzetilmesindedir.

Veri madenciliği veritabanı sistemleri, veri ambarları, yapay öğrenme, istatistik, algoritma, veri görselliği, yapay sinir ağları gibi konuları kapsayan disiplinler arası bir alandır. Bu nedenle yeni bir disiplin olmasına karşın uygulama alanı oldukça geniştir.

Veri madenciliği çoğunlukla veri ambarı ve ÇİHİ gibi diğer veri depolama ve veri değiştirme teknikleri ile ilişkilidir. Bu teknikler çerçevesinde ortak kullanılan bazı temel kavramlar aşağıda açıklanmıştır.

Madencilik modeli, veri madenciliği algoritması tarafından derlenen verilerin fiziksel yapısıdır. Madencilik algoritması, tanımlanmış bir model üzerinde uygulanarak sonuç kümesini oluşturur. (Seidman C., 2001)

Örüntüler, birbirleri arasındaki ilişkiyi ortaya çıkarmak için yeterli sıklıkta tekrar eden veri kümelerinden oluşan olaylar kümesidir. Örneğin, bir market alışverişinde süt satın alımı esnasında aynı zamanda yumurta da alınıyor ve bu işlem yeterli sayıda tekrar ediyorsa, bu durum tahmin edilebilir bir örüntü olarak algılanacaktır.

Durumlar; tarihsel veri üzerindeki her bir öge, veri madenciliği modeli tarafından kullanılan bir durumdur. Market alışverişi örneğinde yer alan her bir satış işlemi bir durumu ifade edecektir.

Veri madenciliği algoritmaları, durumları orijinal veri kaynaklarından veri madenciliği modellerine dönüştüren matematiksel ve istatistiksel algoritmalarıdır.

### 3.1. Veri Madenciliği Yöntembilimi

Veri madenciliği; veri kaynağının seçimi, analiz edilecek kayıtların bazı dönüşümlerle hazırlanması, uygun bir veri madenciliği algoritmasının çalıştırılması, sonuçların değerlendirilmesi süreçlerinden oluşan ve geçerli sonuca ulaşılan kadar kendini tekrarlayan bir süreçtir.

- **Problemin analiz edilmesi;** kaynak veri tabanları incelenerek verilerin madencilik için doğru ve yeterli kriterlere sahip olup olmadıkları incelenir. Verinin uygunluğu, verinin kalitesi ve çokluğu ile doğrudan ilişkilidir. Veri madenciliğinden beklenen sonuç kümesinin doğruluğu için, var olan verilerin doğru türdeki bilgileri içerdiğinden emin olunması gerekir.
- **Verinin çekilmesi ve temizlenmesi;** dış veri kaynaklarına erişilerek bir veri ambarı bünyesinde temizlenmiş geçerli verilerin toplanması sağlanır. Bu adım 'Veri Ambarı' bölümünde 2.2 numaralı başlık altında detaylı olarak açıklanmıştır.
- **Verinin geçerli kılınması;** temizlenerek veri ambarına yüklenen veriler tekrar incelenerek tüm verilerin varlığından ve tamlığından emin olunur.
- **Madencilik modelinin oluşturulması ve eğitilmesi;** uygun madencilik algoritmasının seçilerek modele uygulanması sağlanır. Bu noktada, elde edilen veriler gözden geçirilerek kaynak verilerde yer alan olayların varlığından emin olunmalıdır.
- **Model verisinin sorgulanması;** veri madenciliği ile anlamlı hale getirilmiş veriler karar destek için hazır durumdadır ve yapılacak sorgu sonuçları ile uç kullanıcılara hizmet verirler.
- **Veri madenciliği modelinin geçerliliğinin devamının sağlanması;** madencilik için kullanılan kaynak verilerin yeni gereksinimlere göre değişkenlik gösterebileceği için güncel verilerle modelin de güncel kalması sağlanır.

### 3.2. Veri Madenciliği Modelleri

Veri madenciliğinde kullanılan modeller tahmin edici ve tanımlayıcı olmak üzere iki temel başlık altında toplanır. (Oracle Database Documentation Library,2005)

**Tahmin edici modeller:** Sonuçları bilinen verilerden hareket edilerek bir model geliştirilir ve bu modelden yararlanılarak sonuçları bilinmeyen veri kümeleri için sonuç değerlerinin tahmin edilmesi sağlanır. Bu model kapsamında kullanılan bazı yöntemler aşağıda belirtilmiştir.

- Sınıflandırma (Classification): Nesnelerin ayrı sınıflar içinde gruplanması ve hangi parçanın hangi sınıfa ait olduğunun tahmin edilmesi
- Gerileme (Regression): Yaklaşık ve sürekli değerlerin önceden tahmin edilmesi
- Nitelik Önemi (Attribute Importance): Sonuçları tahmin etmede en etkili niteliklerin bulunması
- Anormallikleri Bulma (Anomaly Detection): Normal verinin özellikleri önemsizmeden nesnelerin teşhis edilmesi

**Tanımlayıcı modeller:** Karar vermeye yardımcı olmak için kullanılacak mevcut verilerdeki örüntülerin tanımlanması sağlanır. Bu model kapsamında kullanılan bazı yöntemler aşağıda belirtilmiştir.

- Kümeleme(Clustering): Veri içindeki doğal grupların bulunması
- İlişkilendirme(Association): Bir veri kümesi içerisindeki ilişkilerin ve gerçekleşme olasılıklarının bulunması
- Özellik Çıkarımı(Feature Extraction): Orijinal özelliklerden yeni özelliklerin ortaya çıkarılması

#### 3.2.1. Sınıflandırma

Sınıflandırma, eldeki verinin parçalara ayrılacak şekilde bölünmesini sağlayacak kuralların bulunmasıyla yapılır. Mevcut verilerden hareket ederek geleceğin tahmin edilmesinde faydalanan ve veri madenciliği teknikleri içerisinde



en yaygın kullanıma sahip olan yöntemdir. Tahmin edilen bağımlı değişken kesikli ve kategorik bir değere sahiptir.

Bir koleksiyonu sınıflandırma, koleksiyonu alt sınıflara ayırma ya da kategorilere ayrılacak şekilde parçalara ayırma şeklinde tanımlanır. (Korth H.F. et al, 1991) Sınıflandırma algoritması, daha önce ayrılmış veriler üzerinde bir model kurar. Bu modeli daha önce yapılan ayrıştırmanın temellerini açıklamak veya yeni verileri sınıflandırmak amacı ile kullanabilir. Bu yolla geçmişe ait verilerden yola çıkarak ileriye dönük tahminler yapmak da mümkün olur.

Sınıflandırma veri üzerinde çalışma ile başlar. Farklı sınıflandırma algoritmaları tahmin edilen nitelik değerleri ile hedef nitelik değerleri arasındaki ilişkiyi bulmak için farklı teknikler kullanır. Bu ilişkiler model için özetlenir ve model yeni veriler üzerine uygulanabilecek hale gelir. Sınıflandırma modeli, test verileri ya da değerlendirme verileri üzerine de uygulanabilir.

Sınıflandırma problemleri ikili ya da çoklu hedefler içerebilir. Sınıflandırma için kullanılan algoritmalar:

- Karar Ağacı Algoritması: Kuralları model şeffaflığını sağladığından genelde iş kullanıcıları, pazarlama analistleri ya da iş analistleri tarafından kullanılır. Şeffaflığın yanı sıra hız ve ölçeklenebilirliği de sağlamaktadır. Veri üzerinde çalışma ve modelleme paralel olarak yürütülebilir. Karar ağacı oluşturulduktan sonra, ağaç kökten yapraklara doğru inilerek kurallar yazılabilir (Mitchell T., 1997). Bu şekilde kural çıkarma veri madenciliği çalışması sonucunun geçerliliğinin test edilmesini sağlar. En yaygın kullanılan algoritmalar 'Classification and Regression Trees(CART), Chi-Squared Automatic Interaction Detector (CHAID), Iterative Dichotomizer Version 3 (ID3) ve C4.5' karar ağacı algoritmalarıdır. Tez kapsamında karar ağacı algoritması kullanılarak veri madenciliği sonuç kümesi üretilmiştir. 3.2.1.1. numaralı bölümde karar ağacı algoritması daha detaylı incelenmiştir.
- Naive Bayes Algoritması: İkili ve çoklu sınıflandırma problemleri için kullanılır. Model oluşturmada ve veri üzerinde uygulamada oldukça hızlıdır.

Ölçeklenebilirliği de oldukça iyidir. Bayes teoremini kullanarak tahminde bulunur.

- Adaptive Bayes Algoritması: Oracle 'a ait bir algoritma olup oldukça hızlıdır, ölçeklenebilir ve hedef özniteliğine göre tahmin edilebilir verilerin çekilmesini sağlayan parametrik olmayan bir yapıya sahiptir. Ayrıca şeffaflığı da sağlar.
- Support Vector Machine Algoritması: Sınıflandırma ve gerileme içeren bir algoritmadır. SVM, gerçek dünya uygulamaları ile verimli çalışır. Text sınıflandırması, el yazısı karakterlerin algılanması, görüntülerin ayırt edilmesi gibi işlemlerde bu algoritma kullanılır.
- Yapay Sinir Ağları Algoritması: Bu yöntemde amaçlanan işlem, birbirine bağlı basit işlemci ünitelerinden oluşan bir ağ üzerine dağıtılmıştır [Bishop, 1996]. Kullanılan öğrenme algoritmaları veriden üniteler arasındaki bağlantı ağırlıklarını hesaplar.

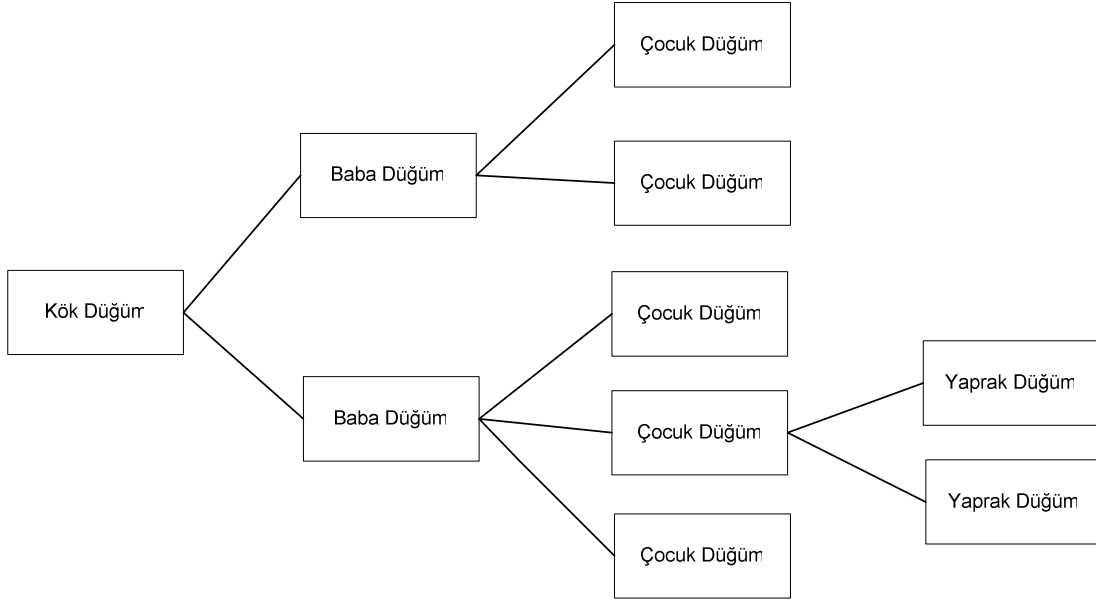
Çizelge 3.1 Sınıflandırma Algoritmaları Karşılaştırması

| Özellik            | Karar Ağacı       | Naive Bayes       | Adaptive Bayes                               | Support Vector Machine          |
|--------------------|-------------------|-------------------|--|---------------------------------|
| Hız                | Hızlı             | Çok Hızlı         | Hızlı  | Hızlı ve Etkin Öğrenme Destekli |
| Doğruluk           | Birçok alanda iyi | Birçok alanda iyi | Birçok alanda iyi                            | Değerli                         |
| Şeffaflık          | Kuralları var     | Kuralları yok     | Sadece tek özellik kurulumunda kuralları var | Kuralları yok                   |
| Eksik Değer Yorumu | Eksik Değer       | Eksik Değer       | Eksik Değer                                  | Seyrek Değer                    |

### 3.2.1.1. Karar Ağacı Algoritması

Karar ağacı, en yaygın kullanılan sınıflandırma algoritmasıdır. Sınıflandırmayı sağlayacak bir dizi kuralın oluşturulması sağlanır. Ağaçta yer alan her bir düğüm,

ilgili düğümün ve o düğümle bağlantılı diğer düğümlerin işlevlerini açıklayan kurallar kümesi olarak ifade edilir. Şekil 3.1 üzerinde karar ağacı veri yapısına ait genel görünüm yer almaktadır. Başlangıç noktasını gösteren kök düğüm üzerinden yaprak düğümlere doğru ilerleyerek, her düğümün temsil ettiği kurallar dizisine göre sonuç kümesi gözlemlenir. Kolay anlaşılır olması tercih edilmesindeki en büyük faktörlerden birisidir.



Şekil 3.1 Genel Karar Ağacı Yapısı Görünümü

Karar ağaçlarına gereksinim duyulabilecek durumları aşağıdaki şekilde toplayabiliriz; (Seidman, C.,2001)

- Potansiyel müşteri grubunu etkileyen bir veri kümesi içerisinde güvenilir bir bölümlenme yapmak istendiğinde
- Bir değişken üzerindeki değişikliğin başka bir değişkeni nasıl etkilediğinin anlaşılabilmesi için değişkenler arasındaki ilişkilerin belirlenmesi istendiğinde
- Veri tabanında bulunan veriler için, ağaç yapısı kullanarak değişkenler arasındaki ilişkilerin görsel bir sunumunun yapılması istendiğinde

- Karmaşık nitelik ve kategoriler için kestirim yapmak için gerekli olanların tespit edilmesi istendiğinde
- Bir veri kümesi içerisinde bulunan ve bir amaç için kullanılabilir önemli değişkenlerin belirlenmesi istendiğinde

Örnek: Sörf yapmak isteyen sporcular için hava ve deniz koşullarını değerlendirerek 'sörf yapılabilir' ya da 'sörf yapılamaz' şeklinde sonuç elde edilmesini sağlayan karar ağacının oluşturulmasını ele alalım.

Çizelge 3.2 Sörf örneği için kural oluşturmak için kullanılacak nitelikler

| Nitelik                | Olası Değerler                |
|------------------------|-------------------------------|
| Dalga Boyu             | Çok Büyük, Büyük, Orta, Küçük |
| Sıcaklık               | Sıcak, Soğuk, Dondurucu       |
| Serinlik               | Çok Serin, Serin Değil, Orta  |
| Dalga Kırılım Kalitesi | İyi, Kötü, Berbat             |

Çizelge 3.3 Sörf Örneği İçin Öğrenme Verileri

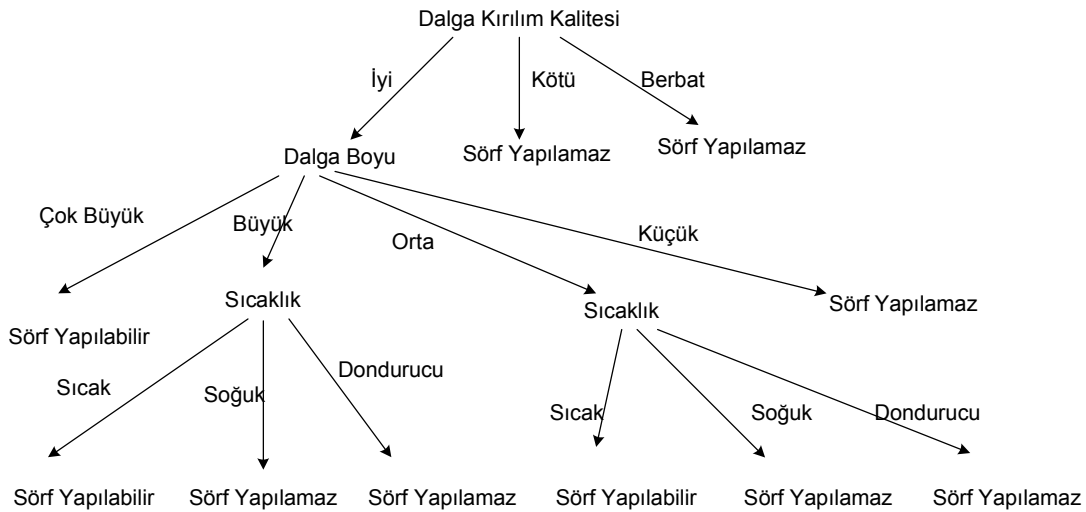
| Dalga Boyu | Sıcaklık  | Serinlik    | Dalga Kırılım Kalitesi | Karar            |
|------------|-----------|-------------|------------------------|------------------|
| Çok Büyük  | Sıcak     | Orta        | İyi                    | Sörf Yapılabilir |
| Büyük      | Sıcak     | Serin Değil | İyi                    | Sörf Yapılabilir |
| Orta       | Sıcak     | Orta        | İyi                    | Sörf Yapılabilir |
| Küçük      | Sıcak     | Çok Serin   | İyi                    | Sörf Yapılamaz   |
| Çok Büyük  | Soğuk     | Orta        | İyi                    | Sörf Yapılabilir |
| Çok Büyük  | Dondurucu | Çok Serin   | İyi                    | Sörf Yapılabilir |
| Çok Büyük  | Soğuk     | Serin       | İyi                    | Sörf Yapılabilir |
| Çok Büyük  | Sıcak     | Serin       | İyi                    | Sörf Yapılabilir |
| Büyük      | Soğuk     | Serin Değil | İyi                    | Sörf Yapılamaz   |
| Büyük      | Dondurucu | Serin       | İyi                    | Sörf Yapılamaz   |
| Büyük      | Soğuk     | Orta        | İyi                    | Sörf Yapılamaz   |
| Büyük      | Soğuk     | Serin       | İyi                    | Sörf Yapılamaz   |
| Büyük      | Sıcak     | Serin       | İyi                    | Sörf Yapılabilir |

Çizelge 3.3 Sörf Örneği İçin Öğrenme Verileri (devam)

| Dalga Boyu | Sıcaklık  | Serinlik    | Dalga Kırılım Kalitesi | Karar            |
|------------|-----------|-------------|------------------------|------------------|
| Büyük      | Sıcak     | Orta        | İyi                    | Sörf Yapılabilir |
| Orta       | Soğuk     | Serin Değil | İyi                    | Sörf Yapılamaz   |
| Orta       | Dondurucu | Serin       | İyi                    | Sörf Yapılamaz   |
| Orta       | Soğuk     | Orta        | İyi                    | Sörf Yapılamaz   |
| Orta       | Soğuk     | Serin       | İyi                    | Sörf Yapılamaz   |
| Orta       | Sıcak     | Serin       | İyi                    | Sörf Yapılabilir |
| Orta       | Sıcak     | Orta        | İyi                    | Sörf Yapılabilir |
| Küçük      | Soğuk     | Çok Serin   | İyi                    | Sörf Yapılamaz   |
| Küçük      | Dondurucu | Serin       | İyi                    | Sörf Yapılamaz   |
| Küçük      | Sıcak     | Orta        | İyi                    | Sörf Yapılamaz   |
| Küçük      | Soğuk     | Orta        | İyi                    | Sörf Yapılamaz   |
| Çok Büyük  | Soğuk     | Serin Değil | Kötü                   | Sörf Yapılamaz   |
| Çok Büyük  | Sıcak     | Serin Değil | Kötü                   | Sörf Yapılamaz   |
| Çok Büyük  | Sıcak     | Orta        | Kötü                   | Sörf Yapılamaz   |
| Çok Büyük  | Soğuk     | Serin       | Kötü                   | Sörf Yapılamaz   |
| Çok Büyük  | Sıcak     | Serin       | Kötü                   | Sörf Yapılamaz   |
| Çok Büyük  | Dondurucu | Serin       | Kötü                   | Sörf Yapılamaz   |
| Çok Büyük  | Dondurucu | Orta        | Kötü                   | Sörf Yapılamaz   |
| Büyük      | Soğuk     | Serin Değil | Kötü                   | Sörf Yapılamaz   |
| Büyük      | Sıcak     | Serin Değil | Kötü                   | Sörf Yapılamaz   |
| Büyük      | Sıcak     | Orta        | Kötü                   | Sörf Yapılamaz   |
| Büyük      | Soğuk     | Serin       | Kötü                   | Sörf Yapılamaz   |
| Büyük      | Sıcak     | Serin       | Kötü                   | Sörf Yapılamaz   |
| Büyük      | Dondurucu | Serin       | Kötü                   | Sörf Yapılamaz   |
| Büyük      | Dondurucu | Orta        | Kötü                   | Sörf Yapılamaz   |
| Orta       | Soğuk     | Serin Değil | Kötü                   | Sörf Yapılamaz   |
| Orta       | Sıcak     | Serin Değil | Kötü                   | Sörf Yapılamaz   |
| Orta       | Sıcak     | Orta        | Kötü                   | Sörf Yapılamaz   |
| Orta       | Soğuk     | Serin       | Kötü                   | Sörf Yapılamaz   |

Çizelge 3.3 Sörf Örneği İçin Öğrenme Verileri (devam)

| Dalga Boyu | Sıcaklık  | Serinlik    | Dalga Kırılım Kalitesi | Karar          |
|------------|-----------|-------------|------------------------|----------------|
| Orta       | Sıcak     | Serin       | Kötü                   | Sörf Yapılamaz |
| Orta       | Dondurucu | Serin       | Kötü                   | Sörf Yapılamaz |
| Orta       | Dondurucu | Orta        | Kötü                   | Sörf Yapılamaz |
| Küçük      | Soğuk     | Serin Değil | Kötü                   | Sörf Yapılamaz |
| Küçük      | Sıcak     | Serin Değil | Kötü                   | Sörf Yapılamaz |
| Küçük      | Sıcak     | Orta        | Kötü                   | Sörf Yapılamaz |
| Küçük      | Sıcak     | Serin       | Kötü                   | Sörf Yapılamaz |
| Küçük      | Dondurucu | Serin       | Kötü                   | Sörf Yapılamaz |



Şekil 3.2 Sörf Örneği İçin Oluşan Karar Ağacı

Çizelge 3.3 kapsamında verilen öğrenme verilerine göre oluşan karar ağacı yapısı Şekil 3.2 ile gösterilmiştir. Karar ağacını incelediğimizde ‘Serinlik’ niteliğinin karar ağacını oluşturan kuralları belirlemede kullanılmadığı, ‘Dalga Kırılım Kalitesi’ niteliğinin ‘Kötü ve Berbat’ durumları için diğer nitelikler önemsenmeden ‘Sörf Yapılamaz’ sonucuna varıldığı görülmektedir. Ağaç üzerinde kökten yapraklara doğru inerek tanımlanmış olan olay kapsamında türetilen kuralları görebiliriz.

Örneğin, dalga kırılım kalitesi iyi, dalga boyu büyük ve sıcaklığın sıcak değerleri için 'Sörf Yapılabilir' sonucu bir kuralı ifade etmektedir.

Sınıflandırma modelleri oluştururken, hem model kurmak için öğrenme verileri hem de ne kadar doğru çalıştığını tespit etmek için test verileri olmalıdır.

### **3.2.2. Gerileme**

Regresyon modelleri, sınıflandırma modelleri ile hemen hemen aynıdır. İkisi arasındaki fark regresyon sayısal ya da sürekli hedef niteliklerle çalışırken; sınıflandırmanın kesikli ya da kategorik hedef nitelikleri üzerinde çalışmasından kaynaklanmaktadır. Eğer hedef nitelik kayan noktalı ya da tamsayı değerler içeriyorsa regresyon tekniği kullanılabilir. Eğer dizgi ya da tamsayı gibi kategorik değerler içeriyorsa sınıflandırma teknikleri kullanılır.

SVM ile hem sınıflandırma hem de regresyon modelleri oluşturulabilmektedir.

### **3.2.3. Nitelik Önemi**

Nitelik önemi, hızı artırmak ve çok sayıda nitelik içeren veri tabloları üzerinden oluşturulmuş sınıflandırma modellerinin doğruluk olasılığını artırmak için kullanılır.

Niteliklerin sayısı arttıkça sınıflandırma modeli oluşturmak için harcanan zaman da artar. Model oluşturma sadece seçilen nitelikler üzerinden yapılır.

Az sayıda nitelik kullanmak çoğunlukla doğrulukta kayba neden olmaz. Çok fazla nitelik kullanmak ise modele etki eder performansını ve doğruluğunu azaltır. En az sayıda nitelik kullanmak etkili bir işletim zamanı sağlayabileceği gibi etkili modeller oluşmasını da sağlar.

Oracle veri madencilik aracı, bu yöntem kapsamında 'Minimum Descriptor Length' algoritmasını kullanır.

### **3.2.4. Anormallikleri Bulma**

Anormalliklerin tespiti, garip ve anormal örneklerden oluşur. Böyle özellikleri teşhis etmek sahtekârlıkları ve internet korsanlığını önlemek gibi problemlerde kullanışlıdır. Anormalliklerin tespiti, verinin veri dağıtımını için tipik olup olmadığına karar verir, dağıtıma uygun olmayan varlıkları belirler.

'One-Class SVM' algoritması bu yöntem için kullanılan bir algoritmadır.

### 3.2.5. Kümeleme

Veriyi tanımlamak için iyi bir yöntemdir. Pek çok durum varsa ve açık doğal bir gruplanma yoksa kümeleme bir veri kümesi içindeki doğal kümeleri bulmanıza yardımcı olur.

Bu yöntem, veri içinde gömülü grupları bulmakla görevlidir. Küme, bazı yönlerden birbirine benzeyen veri nesnelere koleksiyonudur. İyi bir kümeleme fonksiyonu, kaliteli kümeler üretir. Yani birbirine çok benzeyen verileri öbekler.

Bu yöntem ayrıca güdümlü fonksiyonlarda kullanılmak üzere homojen grupların oluşturulmasında da kullanılır. Hedef niteliklere gerek duymaz.

'k-means' ve 'Orthogonal Partitioning Clustering' algoritmaları örnek olarak verilebilir.

### 3.2.6. İlişkilendirme

Genelde bir nesne kümesi üzerindeki ilişkileri ve alt ilişkileri araştırmak için kullanılır. Pazarlama sepeti analizleri, doğrudan pazarlama, katalog tasarımı ve diğer iş karar-görev işletme gibi durumlarda veri analizi için kullanılır.

İlişkilendirme modelleri, geniş müşteri işlemlerinde uygun olayları ve nesnelere ortaya çıkarabilirler. 'Barkod' un kullanılmaya başlanmasıyla satışı yapılan ürünler hakkında verileri toplamak ve saklamak oldukça kolay hale gelmiştir. İlişkilendirme modelleri genelde bu tür satış verileri için kullanılır. Ancak pek çok farklı uygulama için kullanılması da olasıdır. İlişkilendirme kurallarını bulmak promosyonlarda ve pazarlama işlemlerinde değerli, işe yarar olmakla birlikte katalog tasarımı, mağazanın ürün yerleşimi, müşteri kesimleri, web sayfalarının kişiselleştirilmesi gibi konularda da kullanılmaktadır.

Genel olarak, ilişkilendirme modelleri müşteri isteklerini analiz ederek iş trendlerini keşfetmek için kullanılır. Bu web sayfaları için de uygulanabilir. Örneğin şirket X'in şöyle bir birleştirme kuralı olsun: A ve B %80 güvenle C'yi gösterir. A, B ve C 'nin web sayfaları ulaşımı olduğunu düşünelim. A ve B sayfalarını ziyaret eden bir kullanıcı ister doğrudan bir bağlantı olsun ister olmasın C sayfasını %80 olasılıkla



ziyaret edecektir. Bu bilgi A ve B sayfalarından C sayfasına doğrudan geçişin konulması için kullanılabilir. Bu gibi bilgiler web sunucuları için önemli bilgilerdir.

İlişkilendirme modellerinin hesaplanabilecek pek çok özelliği vardır. Kurallarla ilişkili olarak destek ve uyum özellikleri hesaplanır. Destek değeri ile kuralın hem öncel hem de sonuç kesimlerinin nüfusun yüzde kaçı tarafından sağlandığı gösterilir. Uyum değeri ile kuralın öncel kesiminin sağlandığı durumların yüzde kaçında kuralın sonuç kesiminin de sağlandığı gösterilir.

İlişkilendirme yöntemi için 'Apriori' algoritması kullanılır.

### **3.2.7. Özellik Çıkarımı**

Orijinal veriyi kullanarak bir özellik kümesi yaratabilir. Özellik, verinin ilginç karakteristiklerini içeren niteliklerin bir bileşimidir. Özellik çıkarımının bazı uygulamaları, gizli mantık analizleri, veri sıkıştırma, veriyi sıkışmış halinden kurtarma, özellik fark etme olarak tanımlanabilir. Bu güdümlü fonksiyonların hızını ve etkililiğini artırmak için de kullanılabilir. Özellik çıkarma belli anahtar gruplarıyla gösterilmiş ya da bunların birleşimiyle olan dokümanların temalarını ortaya çıkarabilir. Her tema, özellik, anahtar kelimelerin birleşimi ile gösterilebilir. Dokümanlar sonraları bulunan bu temalarla ifade edilebilir.

Özellik çıkarımı yönteminde 'Non-negative Matrix nFactorization' algoritması kullanılır.

### **3.3. Veri Madenciliği Veri Depoları**

Veri madenciliği için oluşturulan modellerin eriştiği veri depoları veri ambarı olarak tasarlanmış ilişkisel veri tabanlarında ya da ÇİAİ veri yapılarında tutulurlar. Veri deposunun seçimi, veri madenciliği modelini kapsayacak problemin yapısına göre değişkenlik gösterir.

#### **3.3.1. İlişkisel Veri Ambarı**

İlişkisel veri ambarları çoğunlukla ÇİHİ veri tabanı ile aynı ilişkisel motoru kullanır. Fakat her iki sisteme ait veriler birbirinden bağımsız fiziksel sunucularda ya da veri tabanlarında tutulur. Herhangi bir sorgu cümleciğine ait sonuç kümesinin üretilmesi için İVTYS 'ne ait sorgulama kabiliyetleri kullanılır. İlişkisel veri ambarının veri

madenciliği için kaynak olarak seçilmesinin avantajları aşağıda maddeler halinde sunulmuştur.

### **3.3.1.1. Esneklik**

ÇİHİ verileri ile veri ambarının aynı veri motoruna sahip olması durumunda veri aktarım işlemleri İVTYS uygulamaları ile hızlı ve etkin iletişim yapıları kurulur. Özellikle veri motorlarının aynı ürün ve sürümüne sahip olması durumunda veri aktarımları etkin ve esnek olur. (Seidman C., 2001) Veri depolama birimi olarak ÇİAİ yapıları tercih edildiğinde, hem ilişkisel hem de ÇİAİ yapılarının işleyişinde özel çaba harcanmalıdır ve küplerin ÇİHİ sisteminde bulunan verilerin durumlarından doğru olarak etkilendiğinden emin olunmalıdır. Veri aktarım aşamasında hangi kayıtların ilişkisel veri tabanında ve küp yapılarında işletildiğini hangilerinin de işletim için beklediğinin doğru tespit edilmesi ve anlaşılması gerekir.

### **3.3.1.2. Standartlık**

Veri tabanı tasarımcıları ve sistem yöneticileri için ilişkisel veri tabanlarının veri madenciliği modelleri oluşturmak için kullanımını daha kolaydır. Çünkü ÇİHİ sistemleri ile aynı standartları kullanırlar. Veri ambarları kolaylıkla oluşturulup yönetilebilir. Çünkü kullanılan kurallar ve terimler ilişkisel veri tabanı kurallarıdır. Hem ÇİHİ hem de veri ambarı yapıları kolaylıkla doğrudan karşılaştırılabilir. Veri tabanı, veri madenciliği işlemleri için yararlı olan izin kullanıma izin vermektedir. ÇİAİ ile küp veri yapılarının kullanılması durumunda yeni teknikler ve kuralların uygulanması gerekliliği ortaya çıkmaktadır. Verilerin doğruluğunun test edilmesi ya da ÇİHİ verileri ile karşılaştırılmasının yapılması daha zor ve karmaşık işlemler gerektirir.

### **3.3.1.3. En Alt Seviyelerdeki Veri Taneciklerine Kolay Erişim**

Veri madenciliği işleminin gücü her detaydaki veriyi kullanabilmesidir. İlişkisel veri tabanında her düzeyde veri tutulabilirken, ÇİAİ yönteminde detaylı veriler anlamlı gruplar halinde özetlenip kümelendiği için en alt seviyelerdeki verinin tutulması sağlanmamaktadır. Bu nedenle, veri madenciliğinde veri kaynağının ilişkisel veritabanı olarak seçilmesi durumunda çok daha detaylı verinin kullanımı söz konusu olmaktadır. Veri madenciliği modeli, bireysel kayıt detayları ile ilgilenen

'Satış' gibi işlemlerde ÇİAİ uygun değildir. Eğer bir pazarlama bölümü belirli bir tür ürün için kişisel satışlara ait özellik ve durumları önemsemeden satış potansiyelini veri madenciliği ile incelemek istiyorsa, ÇİAİ yapıları çok uygun olmaktadır. Bu durumda, her ürün için kişisel satışlar değil, bir ürün kapsamındaki toplam satış değerleri incelenmektedir. Fakat, bir ürüne ait satışlarda ürünü satın alan müşteri profillerine göre bir veri madenciliği modeli geliştirilmek isteniyorsa ÇİAİ yaklaşımı uygun olmamaktadır. ÇİAİ yapıları, ihtiyaç duyulan girdileri sağlamakta yetersiz kalabilir.

### **3.3.2. ÇİAİ Küpleri**

Geleneksel düz tablo yapılarının aksine, ÇİAİ kapsamında küplere ait boyutlar veri kaynağı olarak kullanılır. Veri madenciliği servisi küpleri oluşturan boyutları kullanarak, madencilik modeli için gerekli olan durumları tespit eder. ÇİAİ küpleri, toplamlar, ortalamalar, standart sapmalar, derecelendirme, kayıt toplamları, matematiksel fonksiyonlar ve finansal fonksiyonları içeren türetilmiş kümelerin anlaşılabilirliğinin artırılması için hazırlanmış yapılardır. ÇİAİ küplerinin veri madenciliği için kaynak olarak seçilmesinin avantajları aşağıda maddeler halinde sunulmuştur.

#### **3.3.2.1. Hız**

ÇİAİ küplerinde, veriler önceden hesaplanmış ve hücreler içerisinde saklanmıştır. ÇİAİ verisi ilişkisel veriden oldukça farklıdır. ÇİAİ küp verisinin hazırlanması oldukça uzun zaman alır. Fakat, boyutlar arasında toplamlar ve ortalamalar üzerinden dolaşım yapıldığında oldukça hızlı bir şekilde yanıt verir. ÇİAİ küplerini besleyen veri ambarı çok büyük yoğunlukta veri taşıyorsa ve veri madenciliği modeli için ÇİAİ küplerinde yeralan veriler yeterli derinliğe ve detaya sahipse, veri madenciliği modelinin ÇİAİ küpleri üzerinde oluşturulması tavsiye edilir. Bu sayede, çok yoğun verinin bulunduğu sistemlerde yanıt alma süresi oldukça azalacaktır.

#### **3.3.2.2. Kolay Dolaşım**

Klasik ilişkisel tablolara nazaran ÇİAİ küplerinde küp üzerinde gezinerek veriyi algılama ve inceleme oldukça kolaydır. Boyutlar üzerindeki seviyelere göre ihtiyaç duyulan kümeleme ve toplama işlemleri kolaylıkla tespit edilerek uygulanabilir.

Basit veri madenciliği projelerinde ilişkisel tablo yapıları uygunken, çok daha karmaşık projelerde düz tabloların kullanımı birçok problem yaratmaktadır. Düz tabloların analiz edilmesi zordur. Çok alt seviyeli bilgi tanecikleri içerir ve karmaşık projelerde kullanıldıklarında verinin bütünlüğünün sağlanması için çok fazla kuralın tanımlanması gerekmektedir. Diğer taraftan ÇİAİ üzerindeki veri madenciliği sistemi çoklu boyutlar üzerinden veriye erişerek bu boyutlar üzerindeki örüntülerin tüm avantajlarını kullanırlar. Örüntüler birleştirilebilir, kararlar çizilebilir ve tahmin edilen sonuç değerleri özetlenmiş veri grubu üzerinden yapılabilir. Dolayısıyla çok yoğun verinin ve kural kümesinin yer aldığı veri madenciliği uygulamalarında, ÇİAİ küpleri üzerinden sistem işletildiğinde boyut üzerinde yapılacak kısıtlar ile daraltılan veri kümeleri için ayrı ayrı örüntüler tanımlanıp, sistem karmaşıklığı azaltılabilir.

### **3.3.2.3. Yüksek Yapısallıkta Boyutlar**

Tutarlı kümeleme ve toplama işlemlerinin yapılabilmesi için boyut ve ölçüm değerlerinin titizlikle tanımlanması gerekir. ÇİAİ küplerinin tasarlanmasında kullanılan yazılım araçları, yüksek yapısallıkta boyutlar tanımlanmasını sağlar. Boyutlar üzerinde yanlış bir sıradüzen tanımlanmak istendiğinde bu konuda kullanıcı uyarılır ya da yapılmak istenilen işlem iptal edilir. Dolayısıyla tanımlanmış olan küpler doğru sıradüzen yapılarına sahip boyutlardan oluşur. Boyutlar üzerindeki toplama ve kümeleme değerlerinin yanıltıcı ya da yanlış verileri içermesi engellenmiş olur. Bu sayede, veri madenciliği modellerinde yanlış tahminler oluşması engellenir.

### **3.3.2.4. Madencilik Sonuçlarının Küp Yapılarında Saklanması**

Microsoft Analysis Services ürünü ile, ÇİAİ küpleri kullanılarak hazırlanan veri madenciliği modellerine ait sonuç kümesi sanal boyutlar ve sanal küp tanımları olarak saklanabilmektedir. Dolayısıyla, veri madenciliği sonuçları üzerinde 'drill-down, roll-up, slice, dice, pivot' gibi standart küp işlemleri yapılarak verinin incelenmesi sağlanabilmektedir. İlişkisel tablolar üzerinden oluşturulan veri madenciliği modelleri, küp yapıları ile ifade edilememektedir.

#### **4. Veri Ambarı ve OLAP Teknolojilerinden Yararlanılarak Karar Destek Amaçlı Raporlama Aracı Gerçekleştirimi**

Tez kapsamında, veri ambarı ve ÇİAİ teknolojileri kullanılarak karar destek amaçlı sorgulama ve raporlama imkanı sunun bir araç tasarlanmış ve gerçekleştirilmiştir. Yapılan çalışmaya ait detaylar maddeler halinde açıklanmıştır.

##### **4.1. Tez Kapsamında Kullanılan Ürünler**

Sorgulama ve raporlama aracının gerçekleştirimi öncesinde bu aracı besleyen verilerin oluşturulması ve çok boyutlu veri yapılarında saklanması gerekmektedir. Bu kapsamda işletimsel veri tabanı ve veri ambarı amaçlı kullandığım VTYS'ler, ÇİAİ için kullandığım ürün ve uygulama yazılımı geliştirme aracına ait temel bilgiler aşağıdaki bölümlerde yer almaktadır.

###### **4.1.1. Oracle 9.2.01. VTYS**

Oracle VTYS, ilişkisel veri tabanı sistemidir. Verilerin tutulduğu iki boyutlu tablolar, belirlenmiş olan kısıtlara bağlı kalarak diğer tablolar ile kurulmuş ilişkiler üzerinden veri bütünlüğünü ve doğruluğunu sağlamaktadır. Veri tabanı, sistem, kullanıcı ve kontrol bilgilerinin tutulduğu fiziksel ve mantıksal yapıları içerir. Veri tabanını yöneten yazılıma Oracle veri tabanı sunucusu denir. Oracle'ın çalışmasını sağlayan yazılıma ve fiziksel veri tabanına Oracle veritabanı yönetim sistemi denir. [Oracle Database Documentation Library,2005] Oracle veri tabanı Enterprise, Standard ve Personal olmak üzere 3 sürüme sahiptir. Tez kapsamında kullanılan işletimsel veri tabanı, Oracle 9.2.0.1 sürüm numaralı VTYS'dir.

###### **4.1.2. Microsoft SQL Server 2000 VTYS**

Microsoft SQL Server 2000 VTYS, ilişkisel veri tabanı sistemidir. Enterprise ve Development sürümleri mevcuttur. Tez kapsamında oluşturulan veri ambarı, SQL Server 2000 Development sürümü üzerinde yer almaktadır. Veri ambarının oluşturulmasında kullanılan veri dönüştürme servisi, SQL Server 2000 ürününün bir parçasıdır.

###### **4.1.2.1. DTS (Data Transformation Service), Veri Dönüştürme Servisi**

Microsoft , SQL Server 2000 ürün ailesi içerisinde bulunan ‘Veri Dönüştürme Servisi’ kullanıcıya verileri başka sistemden iç ortama alma, dış ortama taşıma ve bir sistemden başka bir sisteme dönüştürme imkanı tanır. İşlemler paketler halinde saklanır. Bir paket birden çok görev ve adımdan oluşabilir. Paket içerisinde,

- Yapısal Sorgulama Dili komutları
- Jscript, VBScript ve PerlScript betik dilleri ile hazırlanmış betikler
- Dış uygulamalara bağlantı
- E-posta iletimi
- Diğer paketlere sonuç iletme ya da diğer paketlerden sonuç alma

işlemleri gerçekleştirilir. Paket içerisindeki işlemler arasındaki akış, işlemin gerçekleşme durumuna göre kontrol altında tutulur. Bunun için üç tane kontrol parametresi vardır.

- On Success – Başarı durumu ; İşletilen adımın başarılı olma durumu
- On Failure – Başarısızlık durumu; İşletilen adımın başarısız olma durumu
- On Completion – Her durumda; İşletilen adım tamamlandıktan sonra otomatik olarak bir sonraki adıma geçiş durumu

Herbir adım sonucuna göre hangi görevin işleme alınması gerektiği belirlenerek iş akışı tanımlanır. Performansı artırmak için öncelik sırası önemli olmayan işlemler paralel olarak işletilecek şekilde ayarlanabilir.

Bu servis tez kapsamında, veri tabanında işletimsel sistemlerle oluşturulmuş verilerin veri ambarı üzerine gerekli düzeltmeler yapıldıktan sonra aktarılmasını sağlamak için kullanılmıştır. Bunun için betik dili kullanılarak işletimsel sistem üzerinde çalışan sorgular oluşturulmuştur. Sorgu sonucunda elde edilen veriler hatalardan ayıklanarak gruplanmış ve veri ambarı üzerinde uygun tablolara aktarılmıştır. Toplu veri aktarımının yapılmasından sonra aynı betikler, her gece saat 23:45 ‘te çalışarak ilgili gün içerisinde yeni oluşmuş kayıtlara erişerek veri

ambarının sürekli güncel kalmasını sağlamaktadır. İlgili paket için 'Schedule Package' işlemi seçilerek, veri ambarını besleyen betiklerin yer aldığı paketin her gün düzenli olarak çalıştırılması sağlanmıştır. Bu bölümü ifade eden ekran görüntüsü Şekil 4.1 ile gösterilmiştir.

The image shows a Windows-style dialog box titled "Edit Recurring Job Schedule". The job name is "Amerikan\_OLAP\_DTS". Under the "Occurs" section, "Daily" is selected, and "Every 1 day(s)" is entered. The "Daily frequency" section has "Occurs once at: 23:45:00" selected. The "Duration" section has "Start date: 04.04.2005" and "End date: 21.10.2005" selected, with "No end date" also selected. Buttons for "OK", "Cancel", and "Help" are at the bottom.

Şekil 4.1 Veri Dönüştürüm Servisi Görev Tanım Arayüzü

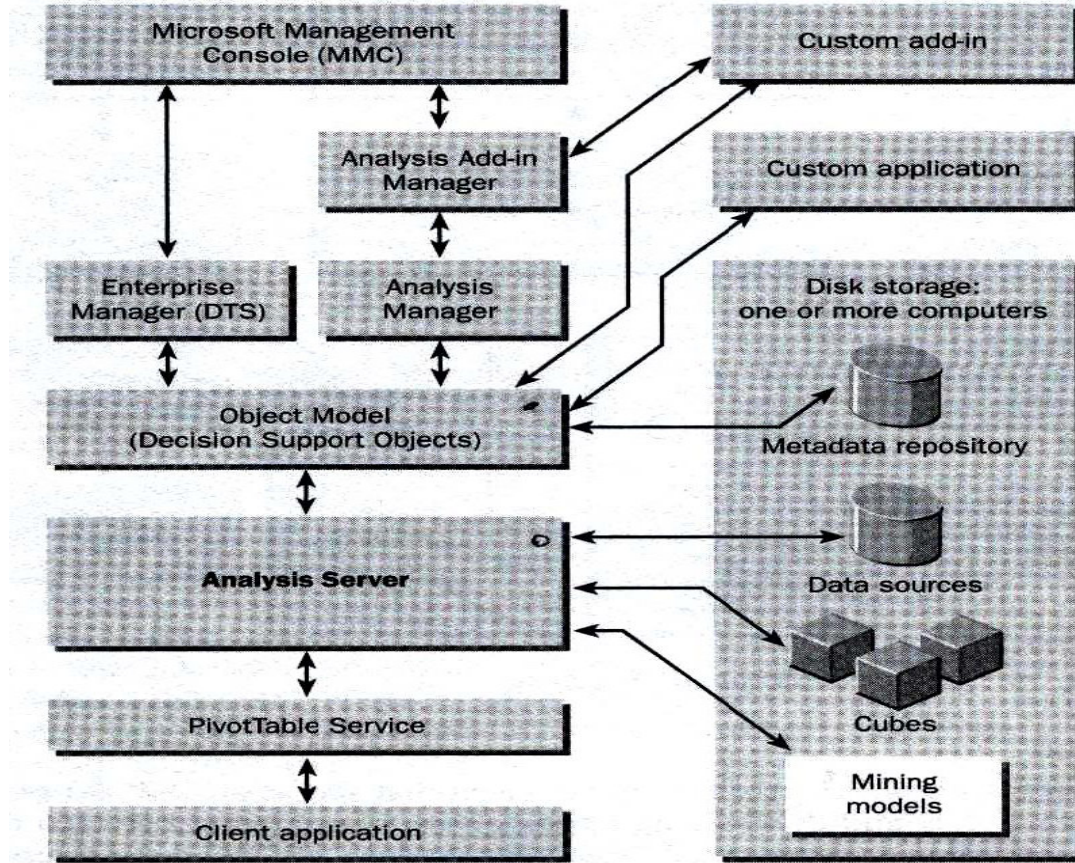
#### 4.1.3. Microsoft Analysis Services

Microsoft SQL Server 2000 Analysis Services, karar destek motorunun ve araçlarının yer aldığı ortamdır [Seidman C., 2001]. Karar destek mekanizmasına ait iki içerik olan veri madenciliği ve ÇİAİ bu ürün kapsamında desteklenmektedir. 'Analysis Services' mimarisi istemci ve sunucu bölümleri olmak üzere ikiye bölünebilir. İstemci bölümü, uç kullanıcılar için arayüz desteğini sağlarken sunucu bölümü, istemci servislere işlevsellik ve güç veren motorların çalışmasını sağlar. İstemci servislere, sunucu bileşenlerinin alt kümesi durumundaki yerel bileşenlerin oluşturulmasında ihtiyaç duyulan işlevselliği sağlarlar. İstemci servislere sunucu bölümünde yer alan ÇİAİ küplerinin oluşturulmasını sağlarken, sunucu bileşenleri bu işlem için gerekli verileri ve görevleri düzenlerler. Microsoft SQL Server 2000

VTYS ile doğrudan bağlantı kurarak çalışabildiği gibi birçok veri tabanı sistemine ODBC üzerinden bağlanabilmektedir.

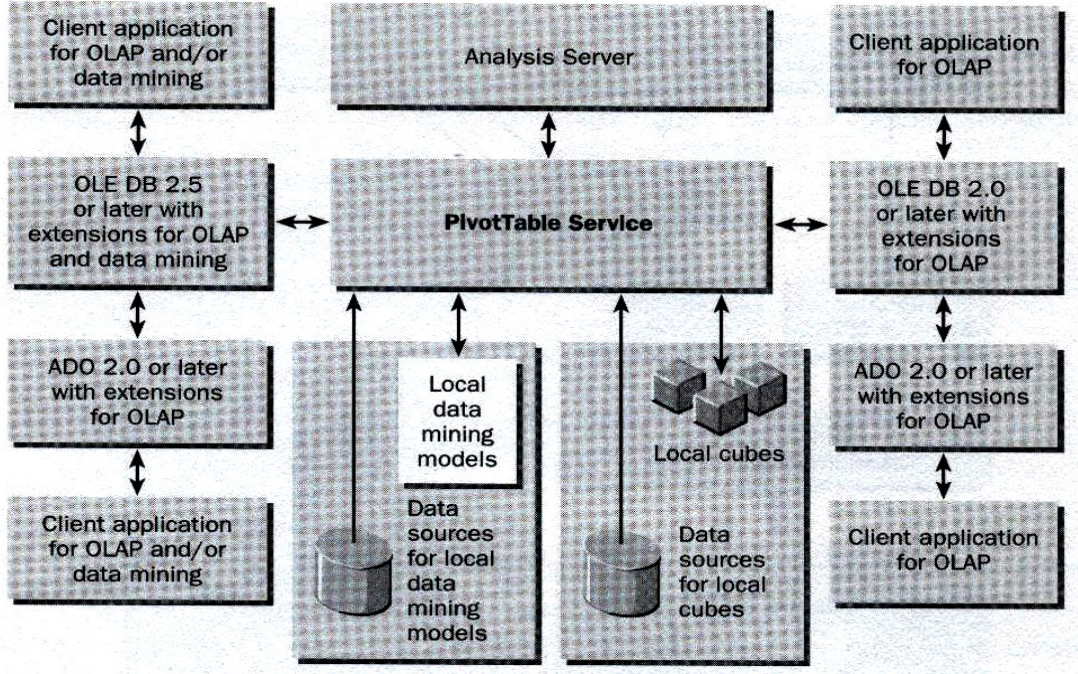
Sunucu ve istemci mimarileri Şekil 4.2 ve Şekil 4.3 ile verilmiştir. İstemci mimarisinde yer alan 'OLE DB for OLAP ve OLE DB for Data Mining' bileşenleri ile ÇİAİ küplerinin ve veri madenciliği çözümlerinin uygulama yazılımları kapsamında ele alınabilmesi sağlanmaktadır.

Ürünün uygulama kolaylığı, anlaşılır olması ve SQL Server 2000 VTYS ile bütünleşik çalışabilmesi tercih edilmesini sağlamıştır.



Şekil 4.2 Analysis Services Sunucu Mimarisi





Şekil 4.3 Analysis Services İstemci Mimarisi

#### 4.1.4. Microsoft Visual Studio .NET 2003

Microsoft firması tarafından sunulan, Windows uygulamaları geliştirilmesini sağlayan yazılım platformudur. .NET çatısı, yazılım geliştiriciler için zengin Windows Forms uygulamalarının geliştirilebilmesi için temiz, nesneye dayalı ve genişletilebilir sınıf kümelerini sunar. Uygulamalar, çok katmanlı dağıtık çözümlerde yerel kullanıcı arayüzü gibi davranır. Microsoft Visual Studio .NET kapsamında, Microsoft Visual Basic .NET, Microsoft Visual C# .NET, Microsoft Visual C++ .NET, Microsoft Visual Java# .NET ve Crystal Reports for Visual Studio .NET yazılım geliştirme araçları yer alır.

#### 4.2. Veri Ambarı ve Veri Küplerinin Tasarımı ve Gerçekleştirimi

Sorgulama ve raporlama aracının gerçekleştirimi öncesinde bu aracın arka planda besleneceği verilerin tutulduğu veri ambarının ve bu ambar üzerinden üretilecek çok boyutlu küp tanımlarının tasarlanması ve oluşturulması gerekmektedir. Bu bağlamda yaptığım çalışmalar aşağıdaki adımlardan oluşmaktadır.

#### **4.2.1. Veri ambarını oluşturacak ÇİHİ tabanlı ilişkisel veritabanının belirlenmesi**

Veri ambarı bir ya da daha fazla işletimsel ya da dış kaynaklı veri kaynaklarından gelen verilerin temizlenerek özetlendiği depo alanıdır. Dolayısıyla, veri ambarı tasarımı öncesinde, veri ambarına veri akışını sağlayacak kaynakların tespit edilmesi gerekir. Tez kapsamında, işletimsel bir sistem olan ve ilişkisel veritabanı üzerinde verilerini saklayan 'Hastane Bilgi Yönetim Sistemi' veri kaynağı olarak seçilmiştir. Bu sistem hastanelerde çalışan uygulama yazılımlarınca girilen verileri içermektedir. Bir hastanın hastaneye gelişinden itibaren gerçekleşen her türlü muayene, tetkik, tıbbi işlem, ameliyat, reçete gibi tüm işlemlerini kapsamaktadır. Tez kapsamında kullandığım veriler, bir 'Hastane Bilgi Yönetim Sistemi' ne ait uygulama yazılımlarınca işlenen örnek verileri içermektedir. Bu veriler Oracle 9.2.0.1 VTYS üzerinde yer almaktadır.

#### **4.2.2. ÇİHİ sisteminin algılanması ve bu sistem üzerinde ihtiyaç duyulabilecek veri ambarı üzerinde bulunması gereken olayların belirlenmesi**

Veri ambarını besleyen veri kaynağının belirlenmesinden sonra, seçilen ÇİHİ sistemine ait veri tabanı tasarımının incelenmesi ve içeriğinin algılanması gerekmektedir. 'Hastane Bilgi Yönetim Sistemi' ne ait iş akışının anlaşılması ve bu iş akışı kapsamında işletimsel sistemin karşılayamadığı ihtiyaçların tespit edilerek bu ihtiyaçların veri ambarı üzerinde karşılanıp karşılanamayacağını belirlenmesi için çalışmalar yaptım. İşletimsel sistemde verilerin hangi detayda ve hangi fiziksel tablolar içerisinde tutulduklarını inceledim. İstatistik raporlarının gerçekleştiriminin oldukça karmaşık olduğunu gözlemledim. Karmaşıklığın hem raporların tasarlanması esnasında hem de verilerin çekilmesi esnasında ortaya çıktığını gördüm. Detay seviyesi oldukça yüksek verilerin tutulması ve veritabanında bulunan veri yoğunluğunun yüksekliği ihtiyaç duyulan istatistiksel verilerin elde edilmesini zorlaştırıyordu. Hem sorgunun karmaşıklığından hem de verilerin çokluğundan dolayı üretilen sorguların cevaplanma süresi oldukça uzun zaman almaktaydı. Bu nedenle, bir istatistik raporunda bulunması gereken sorguların her biri için ayrı raporlar hazırlanarak sorguların geri bildirim süreleri azaltılmaya çalışılmaktaydı. Gözlemlediğim problemlerin veri ambarı ve ÇİAİ yöntemleri ile

aşılabileceğini düşünerek istatistiksel veriler için veri ambarı oluşturmaya karar verdim. Bu kapsamda veri ambarında yer alacak olaylar; ameliyat, muayene, tıbbi işlemler, laboratuvar tetkikleri, patoloji, radyoloji, ölüm, doğum, acil servis işlemleri, reçete ve yatış işlemlerini kapsamaktadır.

#### 4.2.3. Belirlenen olaylar için veri ambarı olarak belirlenmiş ilişkisel veritabanı üzerinde gerekli tablo tanımlarının yapılması

Veri ambarı için oluşturulan ilişkisel veri tabanı Microsoft SQL Server 2000 VTYS üzerindedir. Veri ambarında bulunan ve istatistiksel sorgular hazırlamak amacıyla gerekli verileri tutan tablo tanımları Çizelge 4.1 ve Çizelge 4.2 ile verilmiştir. Olay tablolar, sayılabilir nitelikleri içeren tablolardır ve boyut tabloları ile yabancı anahtarlar üzerinden ilişkilidirler. Veri ambarında bulunan tablolara ait varlık-bağıntı çizenekleri Şekil 4.4 –Şekil 4.9 ile gösterilmiştir.

Çizelge 4.1 Veri Ambarında Bulunan Boyut Tabloları ve Görevleri

| Veri Ambarı Tablosu | Görevi   |
|---------------------|--|
| Acil_Tur            | Acil Serviste yapılan işlem türlerini tutan boyut tablodur.  |
| Ameliyat_Tur        | Ameliyat türlerinin tutulduğu boyut tablosudur.  |
| Anestezi_Durum      | Anestezi durumunu belirten boyut tablosudur. 0 – Yapılmadı, 1- Yapıldı değerlerini tutar.  |
| Bolum_Tip           | Hastanede yer alan bölümlere ait türleri belirten boyut tablodur. Klinik, Poliklinik, Acil Servis, Laboratuvar, Radyoloji, Ameliyathane şeklinde verileri içerir.  |
| Bolumler            | Hastanede yer alan bölümleri tutan boyut tablosudur. Bir bölüme ait bölüm türü Bolum_Tip tablosu ile kurulmuş yabancı anahtara göre belirlenir.  |
| Dogum_Turu          | Doğum türlerini tutan boyut tablosudur. 1 – Normal, 2 – Sezaryen değerlerini tutar.  |
| Doktor              | Hastanede görev yapan doktora ait bilgileri tutan boyut tablosudur. Sicil no, Ad-Soyad ve doktorun uzmanlık alanı saklanır. Uzmanlık alanı Uzmanlik boyut tablosu ile kurulan yabancı anahtar ile tespit edilir. |

Çizelge 4.1 Veri Ambarında Bulunan Boyut Tabloları ve Görevleri (devam)

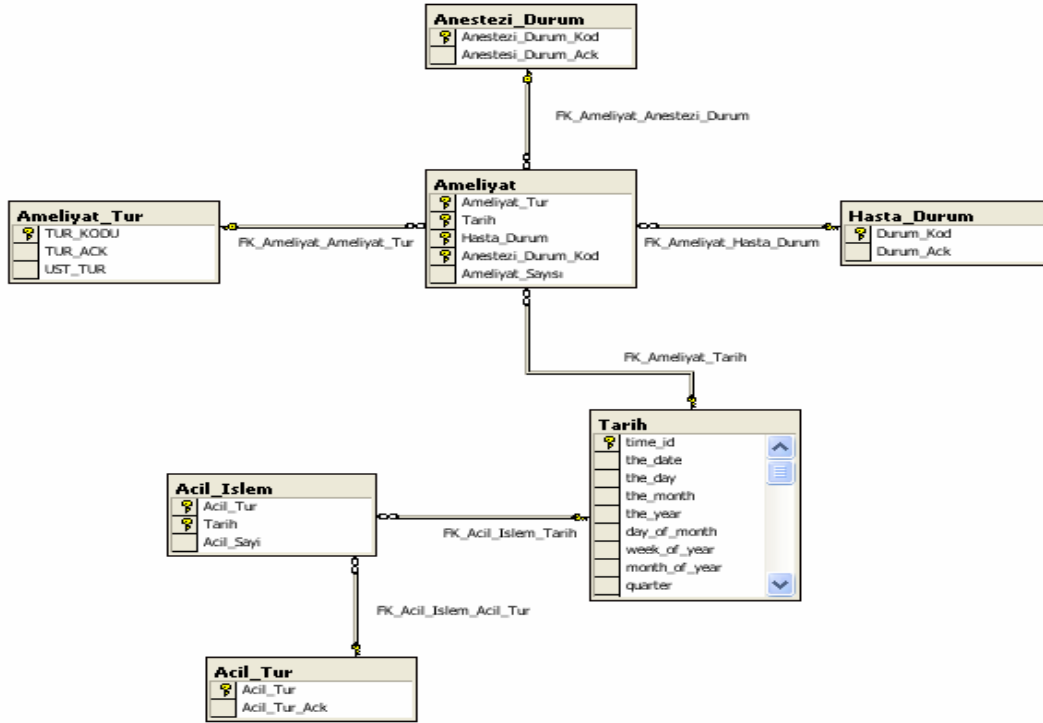
| <b>Veri Ambarı Tablosu</b> | <b>Görevi</b>  |
|----------------------------|--|
| Hasta_Durum                | Hastaneye başvuran hastanın yatan ya da ayaktan hasta olduğunu belirten boyut tablosudur. 0 – Ayaktan , 1- Yatan değerlerini tutar   |
| Islemler                   | Tıbbi işlem türlerini belirten boyut tablosudur. A Grubu Açık Kalp Ameliyatı, Kemik İliği Transplantasyonu, Kanal Tedavisi örnek olarak verilebilir.   |
| Lab_Tetkik_Grup            | Laboratuvar tetkikinin yapıldığı laboratuvar grubunu belirten boyut tablosudur. Mikrobiyoloji, Hematoloji gibi değerleri içerir.   |
| Lab_Tetkik_Tur             | Laboratuvar tetkik türlerini belirten boyut tablosudur. FSH, TSH, LH gibi tetkik türlerini içerir.   |
| Olum_Neden                 | Ölüm nedenlerini belirten boyut tablosudur. Ameliyat Sonucu, Kalp Yetmezliği, Neonatal Bebek Ölümü, Doğum Esnasında Ölüm şeklinde değerlere sahiptir.  |
| Patoloji_Tur               | Patoloji tür bilgisini tutan boyut tablosudur. Akciğer Transbronsial Biopsi , Böbrek Biyopsi gibi veriler içerir.  |
| Patoloji_Step              | Yeni İstek ve Tamamlanmış İstek şeklinde iki değeri içeren boyut tablosudur. Patoloji tetkikinin bulunduğu aşamayı ifade eder.   |
| Radyoloji_Tetkik           | Yeni Doğan Kalça USG, Tüm Abdomen MRG, Water's Grafisi gibi radyoloji tetkik türlerini içeren boyut tablosudur.  |
| Tarih                      | Veri küplerinin tümünde zaman boyutunu ifade etmek için kullanılan boyut tablosudur. Her gün için bir kayıt oluşturulmuştur. Tarih içeriği, haftanın günü, ay açıklaması, yıl, gün, yıl içindeki kaçınıcı hafta olduğu, ay, çeyrek bilgileri tutulur. Örneğin 01.01.2005 tarihi için, '12.01.2005', 'Cumartesi', 'Ocak', '2005', '12', '3', '1', 'Ç1' değerleri bir kayıta saklanmaktadır. |
| Uzmanlik                   | Uzmanlık dallarının tutulduğu boyut tablosudur.  |

Çizelge 4.2 Veri Ambarında Bulunan Olay Tabloları ve Görevleri

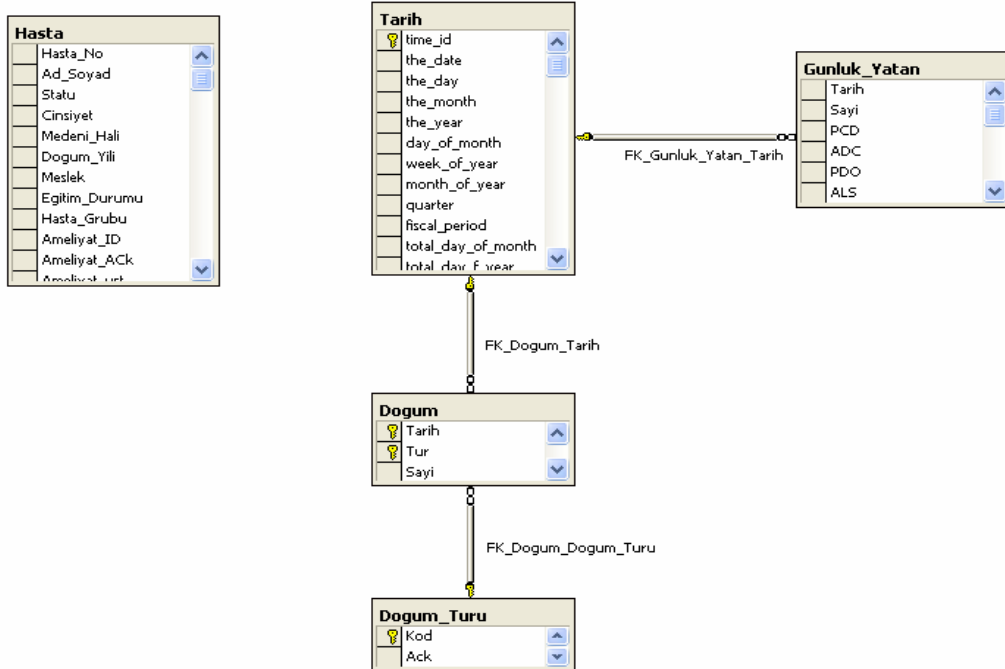
| <b>Veri Ambarı Tablosu</b> | <b>Görevi</b>  |
|----------------------------|--|
| Acil_Islem                 | Acil Serviste yapılan işlemlerin tutulduğu olay tablodur. Acil_Tur ve Tarih boyut tabloları ile ilişkilendirilmiştir.  |
| Ameliyat                   | Anestezi_Durum, Hasta_Durum, Tarih, Ameliyat_Tur boyut tabloları ile ilişkilendirilmiş ve belirli bir tarihte, belirli bir ameliyat türünde, belirli bir hasta ve anestezi durumu için gerçekleşmiş ameliyat sayılarını tutan olay tablodur.   |
| Dogum                      | Dogum_Turu ve Tarih boyut tabloları ile ilişkili olan ve belirli bir tarihte, belirli bir doğum türüne göre gerçekleşmiş doğum sayılarını tutan olay tablodur.   |
| Gunluk_Yatan               | Günlük Yatan Hasta, Ortalama Günlük Yatan, Yatak İşgal Yüzdesi ve Ortalama Yatış Süresi istatistiklerinin hesaplanmasında gerekli olan parametreleri tutan olay tablodur.  |
| Hasta                      | Ameliyat işlemi gerçekleşmiş hastalara ait detaylı bilgileri içeren olay tablodur. Cinsiyet, medeni hali, doğum yılı, statüsü, mesleği, eğitim durumu, hasta grubu ve geçirdiği ameliyat bilgileri tutulur. Bu bilgilere göre veri madenciliği algoritması işletilerek ameliyat türleri ile diğer tutulan özellikler arasında bir ilişki aranmıştır. |
| Lab_Islemler               | Lab_Tetkik_Grup, Lab_Tetkik_Tur, Hasta_Durum ve Tarih boyut tabloları ile ilişki olan olay tablodur. Laboratuar tetkik sayılarının belirtilen boyutlara göre saklanması için tanımlanmıştır.   |
| Muayene                    | Bolum, Tarih, Uzmanlik ve Hasta_Durum boyut tabloları ile ilişki ve muayene sayılarını tutan olay tablodur.  |
| Olum                       | Olum_Neden ve Tarih boyut tabloları ile ilişkili ve ölüm sayılarının nedenler ve tarihlere göre gruplanmasını sağlayan olay tablodur.  |

Çizelge 4.2 Veri Ambarında Bulunan Olay Tabloları ve Görevleri (devam)

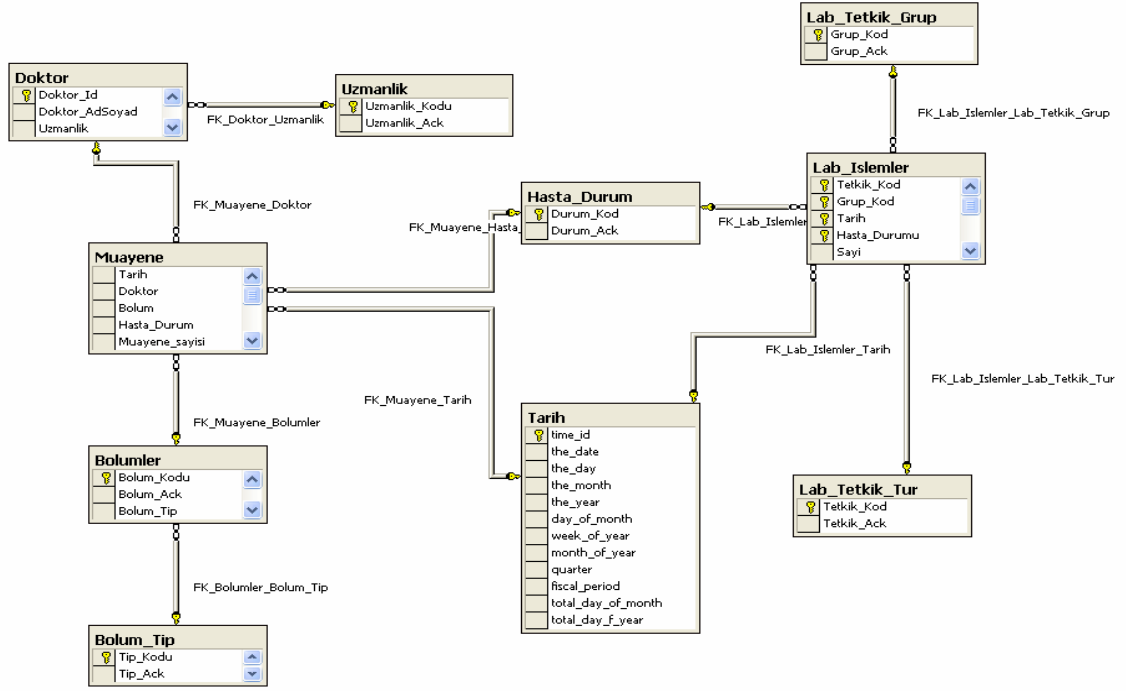
| <b>Veri Ambarı Tablosu</b> | <b>Görevi</b>  |
|----------------------------|--|
| Patoloji                   | Hasta_Durum, Patoloji_Step, Patoloji_Tur ve Tarih boyut tabloları ile ilişkilidir. Patoloji tetkik sayılarının belirtilen boyutlara göre saklanmasını sağlar.                      |
| Radyoloji                  | Radyoloji tetkik sayılarını tutan olay tablodur. Hasta_Durum, Radyoloji_Tetkik ve Tarih boyut tabloları ile ilişkilendirilmiştir.  |
| Recete                     | Belirli bir tarihte yazılmış olan reçete sayılarını belirten olay tablodur.  |
| Tibbi_Islemler             | Yapılmış olan tıbbi işlem sayılarını tutan olay tablodur. Bolum, Hasta_Durum, Islemler ve Tarih boyut tabloları ile ilişkilendirilmiştir   |
| Yatak                      | Hastanede bulunan yatak sayılarının yatak türlerine göre gruplayan olay tablodur. Tarih boyut tablosu ile ilişkilidir.   |
| Yatısı_Yapilan             | Uzmanlik ve Tarih boyut tabloları ile ilişkilidir. Belirli bir tarihte belirli bir uzmanlık dalından yatışı yapılmış olan hasta sayılarını saklamak için kullanılan olay tablodur. |



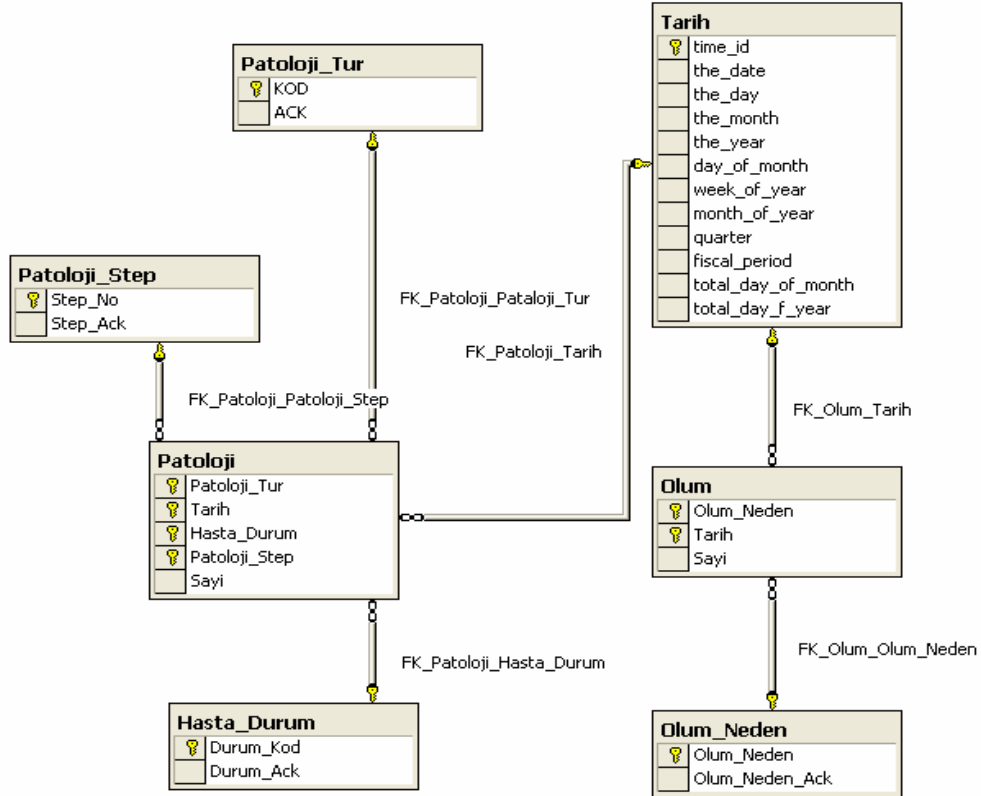
Şekil 4.4 Ameliyat ve Acil İşlem olaylarını gösteren Varlık-Bağıntı Çizeneği



Şekil 4.5 Doğum, Hasta ve Günlük Yatan Hasta olaylarını gösteren Varlık-Bağıntı Çizeneği

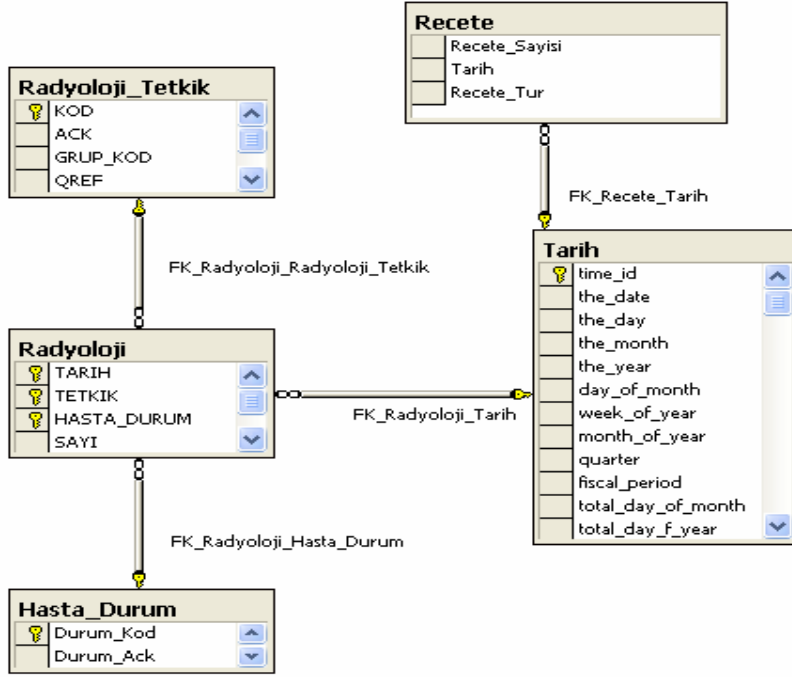


Şekil 4.6 Muayene ve Laboratuvar İşlemlerini gösteren Varlık-Bağıntı Çizeneği

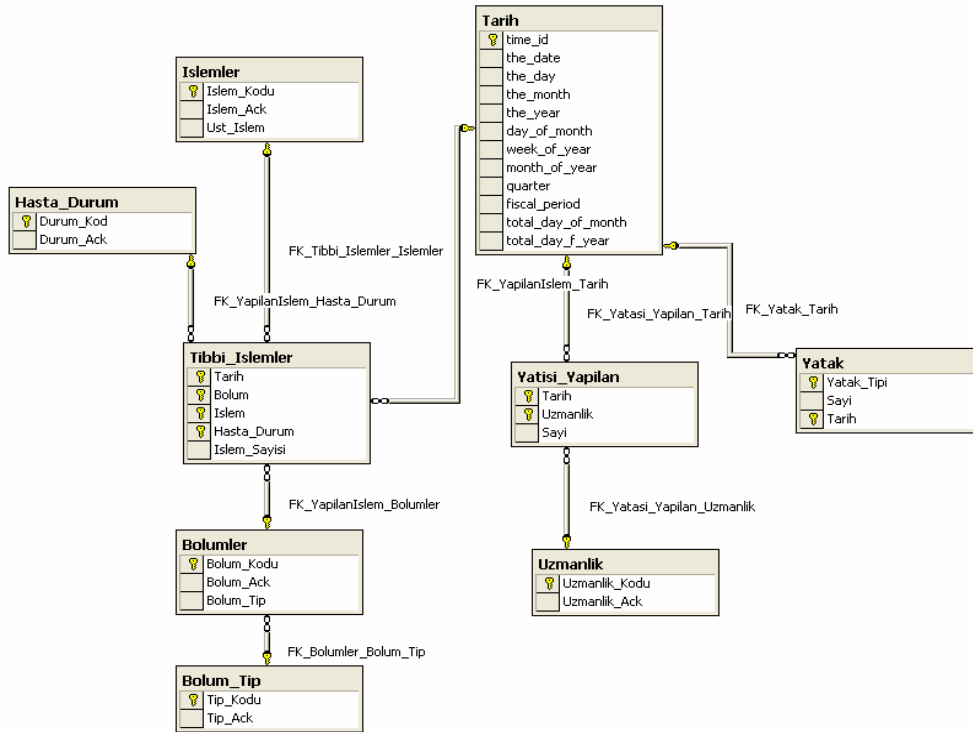


Şekil 4.7 Patoloji İşlemleri ve Ölüm olaylarını gösteren Varlık-Bağıntı Çizeneği





Şekil 4.8 Radyoloji İşlemleri ve Reçete olaylarını gösteren Varlık-Bağıntı Çizeneği



Şekil 4.9 Tıbbi İşlemler, Yatışı Yapılan Hasta ve Yatak olaylarını gösteren Varlık-Bağıntı Çizeneği

#### **4.2.4. Belirlenen olayların beslendiği verilerin temizlenmesi, yeniden biçimlendirilmesi ve veri ambarında uygun tablolara yüklenmesi**

Veri ambarı tasarımı tamamlandıktan sonra hazırlanmış olan tabloların verileri içermesi için gerekli olan verinin çekilmesi, veri dönüştürmesi (temizleme, kaynaştırma, kümeleme, özetleme), verinin yüklenmesi ve verinin güncelliğini koruması için gerekli olan adımlar gerçekleştirilmiştir. Bu işlemlere ait genel açıklamalar 2.2. Veri Ambarının Tasarlanması bölümünde yer almaktadır. Tez kapsamında veri ambarının verilerle yüklenmesi için gereken bölümler Microsoft SQL Server Enterprise Manager ürünü altında yer alan Data Transformation Services (DTS) adlı veri dönüştürme servisi ile yapılmıştır. Betik dili kullanılarak hazırlanan görevler ile işletimsel veritabanında bulunan verinin çekilmesi, temizlenmesi, dönüştürülmesi ve veri ambarına yüklenmesi sağlanmıştır. Oracle VTYS bünyesinde bulunan işletimsel sisteme erişim için ODBC bağlantısı kullanılmıştır. Bağlantı cümlecisi şu şekildedir; "DRIVER={Microsoft ODBC for Oracle}; UID=HBYS; PWD=Admin; SERVER=HBYSserver;"

Bağlantının sağlanabilmesi için Oracle VTYS için tanımlı ODBC sürücüsü kullanılarak, kullanıcı kimliği 'HBYS', kullanıcı şifresi 'Admin' ve bağlantının yapılacağı veri tabanı sunucusu 'HBYSserver' bilgileri ile işletimsel veritabanına erişilmiştir.

Kurulmuş olan bağlantı üzerinde çalıştırılan YSD ile hazırlanmış sorgularla tarihsel bazda istatistiksel verilerin çekilmesi sağlanmıştır. Elde edilen veriler için,

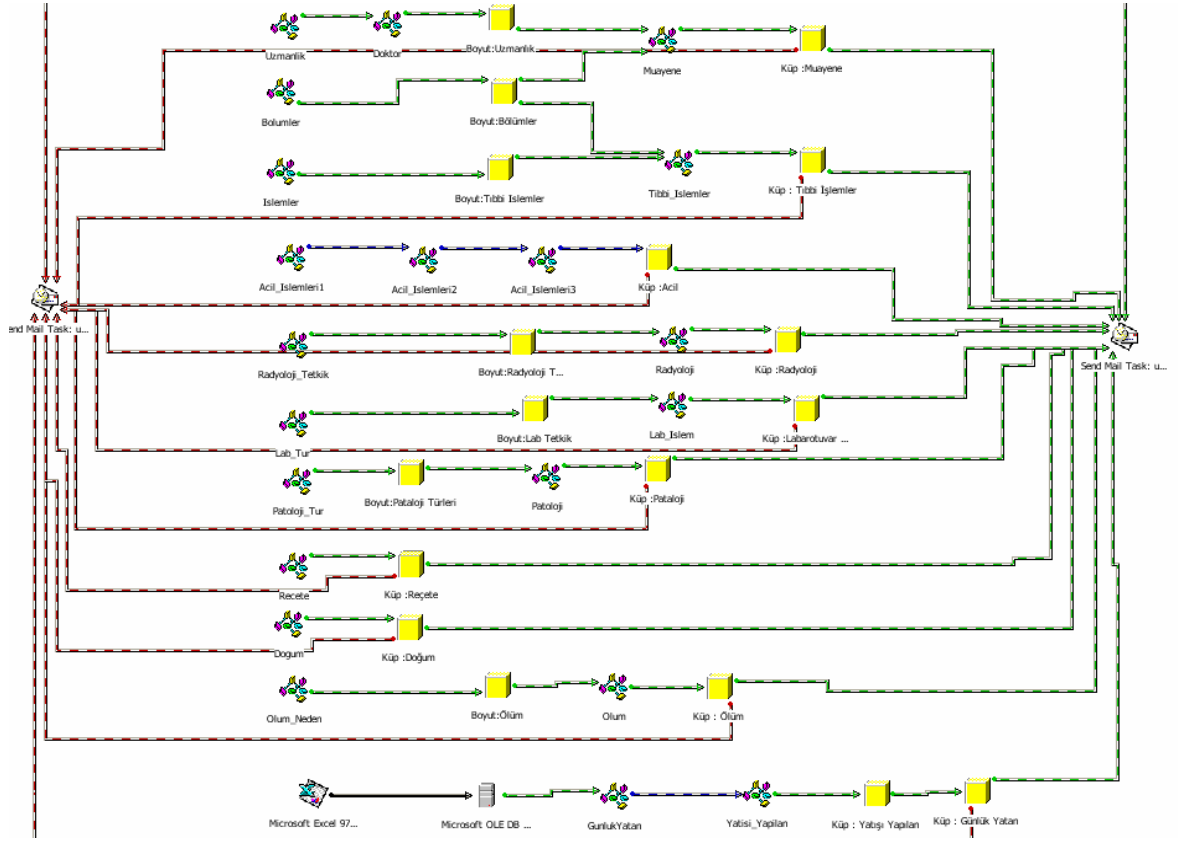
- Bazı sütun içerikleri boş olan kayıtlar varsa, yapılacak işlemin içeriğine göre bu kayıtlar uygun verileri içerecek şekilde doldurulmuştur ya da veri ambarına alınmamıştır.
- Tarih verileri, veri ambarında bulunan 'Tarih' tablosundaki bir kayıta referans verecek şekilde güncellenmiştir. Örneğin 01.02.2004 tarih içeriği 'Tarih' tablosunda 397 kimliğe sahip kayıta karşılık gelmektedir. Çekilen tarihsel veri 397 değeri ile veri ambarına konmuştur.

- Veri ambarı tasarımında kurulmuş olan yabancı anahtarlara uygun olmayan kayıtların aktarımı yapılmamıştır. Örneğin sistemde bulunmayan bir bölümde gerçekleşmiş muayene işlemi ile karşılaşıldığında verinin doğruluğunun bulunmadığı tespit edilerek veri aktarımı yapılmamıştır.
- Hazırlanmış kod tablolarına göre verilerin kodlanmasının sağlanması. Örneğin 'Yapıldı, Yapılmadı' şeklinde alınan anestezi durum bilgilerinin 0 ve 1 kodları ile 'Anestezi\_Durum' kod tablosu ile saklanması sağlanmıştır. Dolayısıyla açıklama olarak gelen veriler kodlanarak veri ambarına aktarılmıştır.
- Yüksek detaya sahip işletimsel veritabanından, veri ambarında tasarlanmış olay tablo yapısına uygun olacak şekilde gruplanmış özet veriler aktarılmıştır. Örneğin, ameliyat işlemleri için veri ambarında belirli bir tarihte, belirli bir ameliyat türünde, belirli bir anestezi ve hasta durumu için gerçekleşmiş ameliyat sayıları tutulmaktadır. Dolayısıyla veriler çekilirken ameliyat türü, tarih, hasta durumu ve anestezi durumuna göre gruplanmış özetlenmiş sayısal içerikler olarak veri ambarına aktarılmıştır.
- Açıklama alanı olarak alınan 'Cinsiyet, Medeni Hali, Meslek, Eğitim Durumu' sütunlarının aynı veriyi ifade edecek şekilde düzenlenmesi sağlanmıştır. Örneğin, Cinsiyet kapsamında aynı sütun için bazı kayıtlarda 'E' bazı kayıtlarda 'Erkek' değeri bulunmaktaydı. Bu örnek için ilgili kayıtlarda 'Erkek' değerinin yer alması sağlanmıştır.

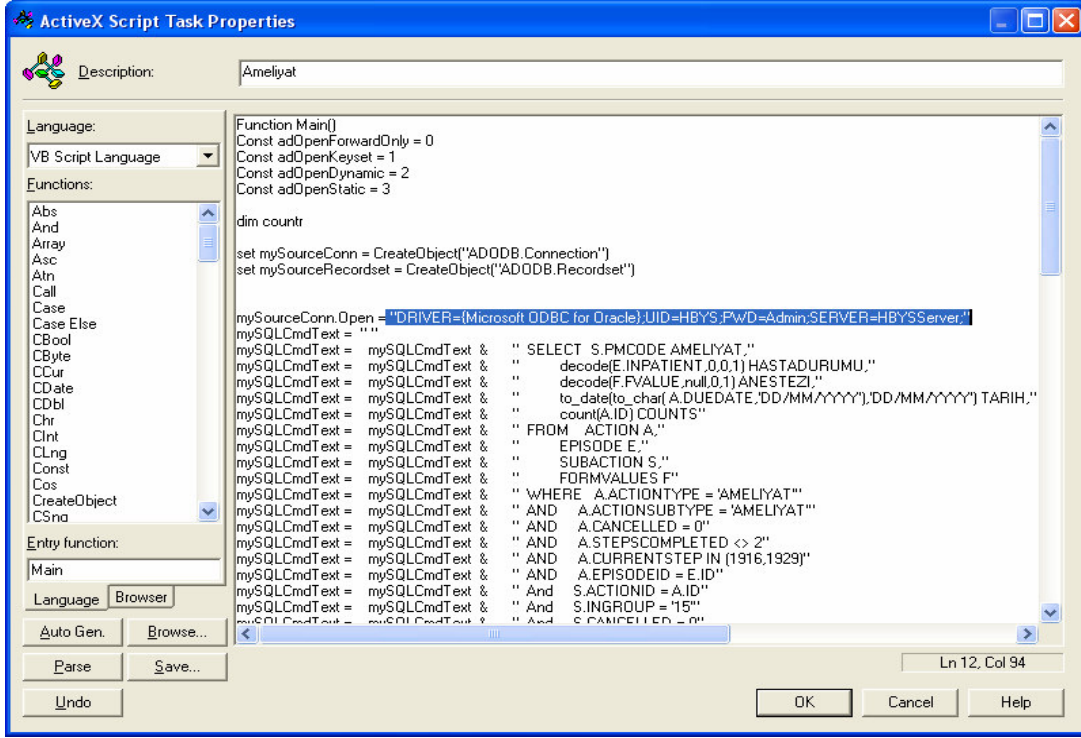
Veri dönüştürme servisi ile tüm boyut ve olay tablolarına ait veri aktarımını sağlayan betikler hazırlanmış ve bu betikler işletilerek verilerin veri ambarına aktarımı sağlanmıştır.

Veri dönüştürme servisinde çalıştırılan pakete ait genel görünüm Şekil 4.10 ile verilmiştir.

Şekil 4.11 ile Ameliyat olay tablosunun yüklenmesinde kullanılan betik dili ile hazırlanmış görev gösterilmiştir.



Şekil 4.10 Veri Dönüştürme Servisi Üzerinde Hazırlanmış Pakete Ait Genel Görünüm



Şekil 4.11 Ameliyat olay tablosunun yüklenmesinde kullanılan betik dili ile hazırlanmış görev içeriği

#### 4.2.5. Oluşturulan veri ambarı üzerinde çok boyutlu veri yapılarının tanımlanması

Verileri de içerecek şekilde oluşturulan veri ambarı üzerinde çok boyutlu küplerin tanımlanması gerçekleştirilmiştir. Bu işlem için Microsoft Analysis Manager aracı kullanılmıştır. Bu araç ile Microsoft SQL Server üzerinde bulunan veri ambarına erişim sağlanarak, boyut ve olay tabloları belirlenerek küp tanımları hazırlanmıştır.

##### 4.2.5.1. Boyutların uygun sıradüzensel yapıda oluşturulması

Her boyut için en az bir tane sıradüzen bulunmalıdır. Aynı boyut için birden çok sıradüzen oluşturulabilir fakat tez kapsamında böyle bir ihtiyaç oluşmadığı için tasarladığım boyutlar tek bir sıradüzenden oluşmaktadır. Çizelge 4.1 ile verilen her boyut tablosu için birer boyut bulunmaktadır. Tasarlanmış boyutlar ve sıradüzen bilgileri Çizelge 4.3 ile gösterilmiştir.

Microsoft Analysis Manager ürününün sağladığı boyut türleri;

- Yıldız şema, tek bir boyut tablosu kullanılarak oluşturulan boyut tanımıdır.
- Kar yağışı şema, birden çok ilişkili boyut tablosu kullanılarak oluşturulan boyut tanımıdır.
- Baba-oğul, tek bir boyut tablosu üzerinde iki ilişkili sütun kullanılarak oluşturulan boyut tanımıdır.
- Sanal boyut, başka bir boyutta yer alan üye tanımından oluşturulan boyut tanımıdır.
- Madencilik modeli, veri madenciliği modelinde oluşturulan tahmin edilebilir sütunları ifade eden boyut tanımıdır.

Çizelge 4.3 Boyut Tanımları

| Boyut Adı    | Kullanılan Boyut Tablosu | Boyut İçeriği  |
|--------------|--------------------------|--|
| Acil_Tur     | Acil_Tur                 | Acil_Tur_Ack sütununa göre acil servis işlem türlerini ifade eden açıklama alanı üzerinde yıldız şema boyutu oluşturulmuştur. Tek seviyelidir.   |
| Ameliyat_Tur | Ameliyat_Tur             | Ameliyat_Tur tablosunda, Tur_Kodu sütunu Ust_Tur sütunuyla kendisine referans vermektedir. Dolayısıyla, aynı tablo üzerinde baba-oğul ilişkisi kurulmuştur. Bu sütunlar üzerinde baba-oğul ilişki kullanılarak oluşturulan boyut tanımıdır. Tabloda yer alan verinin derinliği kadar seviyeye sahiptir. Örneğin, en üst seviyede 'Tüm Ameliyatlar' yer alır. Bir alt seviyede, Genel Cerrahi, bu seviyenin altında Baş ve Boyun yer alır. 4. seviyede de Tiroid Biyopsisi gibi en alt seviyeyi oluşturan ameliyat türleri bulunur. |

Çizelge 4.3 Boyut Tanımları (devam)

| Boyut Adı   | Kullanılan Boyut Tablosu | Boyut İçeriği  |
|-------------|--------------------------|--|
| Anestezi    | Anestezi_Durum           | Anestezi_Durum_Ack sütununa göre anestezi durum açıklamalarının tek seviyeli olarak tutulduğu yıldız şema boyutudur.   |
| Bolum       | Bolumler ,<br>Bolum_Tip  | Bölüm tip açıklaması ve bölüm açıklamasına göre iki seviyeli oluşturulmuş kar yağışı boyutudur. İlk seviyede bölüm tipleri, ikinci seviyede bölümler yer alır.   |
| DogumTuru   | Dogum_Turu               | Doğum türlerinin açıklama alanlarına göre tek seviyeli olarak tutulduğu yıldız şema boyutudur.   |
| Hasta_Durum | Hasta_Durum              | Hasta durum açıklamalarını gösteren Durum_Ack sütunu üzerinden tek seviyeli olarak tanımlanan yıldız şema boyutudur.   |
| Islemler    | Islemler                 | Islemler tablosunda yer alan, Islem_Kodu ve Ust_Islem sütunları birbirine referans vermektedir. Baba-oğul ilişkisine göre tanımlanmış boyutta boyut seviyesi bazı dallarda 3 bazılarında 4 olmaktadır. Bu durum verinin dağılımına göre değişiklik göstermektedir. Örneğin, Tüm İşlemler – Dializ – Hemodializ – Hemofiltrasyon şeklinde 4 seviye olabilirken, Tüm İşlemler – Acil Klinik – Alçı Çıkarma şeklinde 3 seviyede de kalabilir. |
| LabGrup     | Lab_Tetkik_Grup          | Laboratuar tetkik gruplarına ait açıklamalara göre oluşturulmuş tek seviyeli yıldız şema boyutudur.  |
| LabTetkik   | Lab_Tetkik_Tur           | Laboratuar tetkik türlerine göre tek seviyeli yıldız şema boyutudur.   |

Çizelge 4.3 Boyut Tanımları (devam)

| <b>Boyut Adı</b> | <b>Kullanılan Boyut Tablosu</b> | <b>Boyut İçeriği</b>  |
|------------------|---------------------------------|---|
| Patoloji_Step    | Patoloji_Step                   | Patoloji işlemlerinin bulunduğu adımı belirtmek için tutulan açıklama alanı için oluşturulmuş tek seviyeli yıldız şema boyutudur.   |
| Patoloji_Tur     | Patoloji_Tur                    | Patoloji işlem türlerinin açıklama alanına göre tek seviyeli olarak oluşturulduğu yıldız şema boyutudur.  |
| Radyoloji_Tetkik | Radyoloji_Tetkik                | Radyoloji_Tetkik tablosunda yer alan Kod ve Grup_Kod sütunları baba-oğul ilişkisine sahiptir. Dolayısıyla aynı tabloda hem baba hem de çocuk kayıtlar yer alır. Oluşturulan boyut baba-oğul ilişkisine göre kurulmuştur. Örneğin Tüm Radyoloji Tetkikler – MR – Boyun MR şeklinde üç seviye oluşmaktadır. |
| Tarih            | Tarih                           | Yıl ve ay sıradüzenine sahip zaman boyutudur.   |
| Uzmanlik         | Doktor, Uzmanlik                | Uzmanlık açıklaması ve doktor ad, soyad bilgilerine göre hazırlanmış kar yağışı şema boyutudur. İlk seviyede uzmanlık türleri, ikinci seviyede her bir uzmanlık türüne göre doktor ad, soyad bilgileri yer almaktadır.  |
| Uzmanlik_Genel   | Uzmanlik                        | Uzmanlık açıklamasına göre uzmanlık dallarının tek seviyeli olarak tutulduğu yıldız şema boyutudur.   |
| Yatak            | Yatak                           | Yatak tablosunda yer alan Yatak_Tipi sütununa göre hazırlanmış yıldız şema boyutudur. Yatak tip açıklamaları tek seviyeli olarak gösterilmektedir.  |



#### 4.2.5.2. Küp Tanımlarının Yapılması

Tasarlanmış olan boyut tanımlarına göre sayısal ölçülebilir nitelikleri barındıran küpler tasarlanmıştır. Küp tanımları Çizelge 4.4 üzerinde açıklanmıştır.

Çizelge 4.4 Küp Tanımları

| <b>Küp Adı</b> | <b>Kullandığı Boyutlar</b>   | <b>Kullanılan Olay Tablo</b> | <b>Ölçülebilir Değerler</b>   |
|----------------|--|------------------------------|---|
| Acil_Islem     | Tarih, Acil_Tur  | Acil_Islem                   | Acil İşlem Sayısı   |
| Ameliyat       | Ameliyat_Tur,<br>Anestezi, Tarih,<br>Hasta_Durum                                       | Ameliyat                     | Ameliyat Sayısı   |
| Dogum          | Tarih, DogumTuru   | Dogum                        | Doğum sayısı  |
| Gunluk_Yatan   | Tarih  | Gunluk_Yatan                 | Günlük Yatan Hasta Sayısı, Ortalama Günlük Yatan Hasta Sayısı, Yatak İşgal Yüzdesi, Ortalama Yatış Süresi |
| HastaBlg       | Ameliyat_Tur,<br>Cinsiyet,<br>Medeni_Hal,<br>Egitim_Durumu,<br>Meslek,<br>Hasta_Numara | Hasta                        | Ameliyat Sayısı   |
| Lab_Islemleri  | Tarih, LabGrup,<br>LabTetik,<br>Hasta_Durum  | Lab_Islemler                 | Laboratuvar İşlem Sayısı  |
| Muayene        | Tarih, Bolum,<br>Uzmanlik,<br>Hasta_Durum  | Muayene                      | Muayene Sayısı  |
| Olum           | Olum_Neden, Tarih  | Olum                         | Ölen Hasta Sayısı   |

Çizelge 4.4 Küp Tanımları (devam)

| <b>Küp Adı</b>  | <b>Kullandığı Boyutlar</b>                               | <b>Kullanılan Olay Tablo</b> | <b>Ölçülebilir Değerler</b> |
|-----------------|--|------------------------------|-----------------------------|
| Patoloji        | Patoloji_Tur,<br>Patoloji_Step,<br>Hasta_Durum,<br>Tarih | Patoloji                     | Patoloji Sayısı             |
| Radyoloji       | Radyoloji_Tetkik,<br>Hasta_Durum,<br>Tarih               | Radyoloji                    | Radyoloji İşlem Sayısı      |
| Recete          | Tarih  | Recete                       | Reçete Sayısı               |
| Tibbi_Islemler  | Islemler,<br>Hasta_Durum,<br>Bolum, Tarih                | Tibbi_Islemler               | İşlem Sayısı                |
| Yatak           | Yatak, Tarih   | Yatak                        | Yatak Sayısı                |
| Yatisci_Yapilan | Tarih,<br>Uzmanlik_Genel                                 | Yatisci_Yapilan              | Yatış Sayısı                |

Çizelge 4.5 Küp Açıklamaları

| <b>Küp Adı</b> | <b>Açıklama</b>  |
|----------------|--|
| Acil_Islem     | Acil serviste yapılan hasta giriş sayılarının tarih ve işlem türüne göre gruplayan küp tanımıdır.                      |
| Ameliyat       | Ameliyat sayılarının anestezi durum, hasta durum, ameliyat tür ve zaman boyutuna göre gruplayan küp tanımıdır.         |
| Dogum          | Normal ve sezaryen şeklinde tutulan doğum türlerine göre tarihe bağlı olarak doğum sayılarını tutan küp tanımıdır.     |
| Gunluk_Yatan   | Yatan hasta sayılarına ait zaman boyutlu aritmetik ifadelerle oluşturulmuş özet verileri tutan küp tanımıdır.          |
| HastaBlg       | Veri madenciliği kapsamında kullanılan ve hasta detay bilgilerine göre ameliyat bilgilerini tutan küp tanımıdır.       |
| Lab_Islemleri  | Laboratuvar işlem sayılarının tetkik türleri ve grupları, hasta durumu ve zaman boyutuna göre gruplayan küp tanımıdır. |

Çizelge 4.5 Küp Açıklamaları (devam)

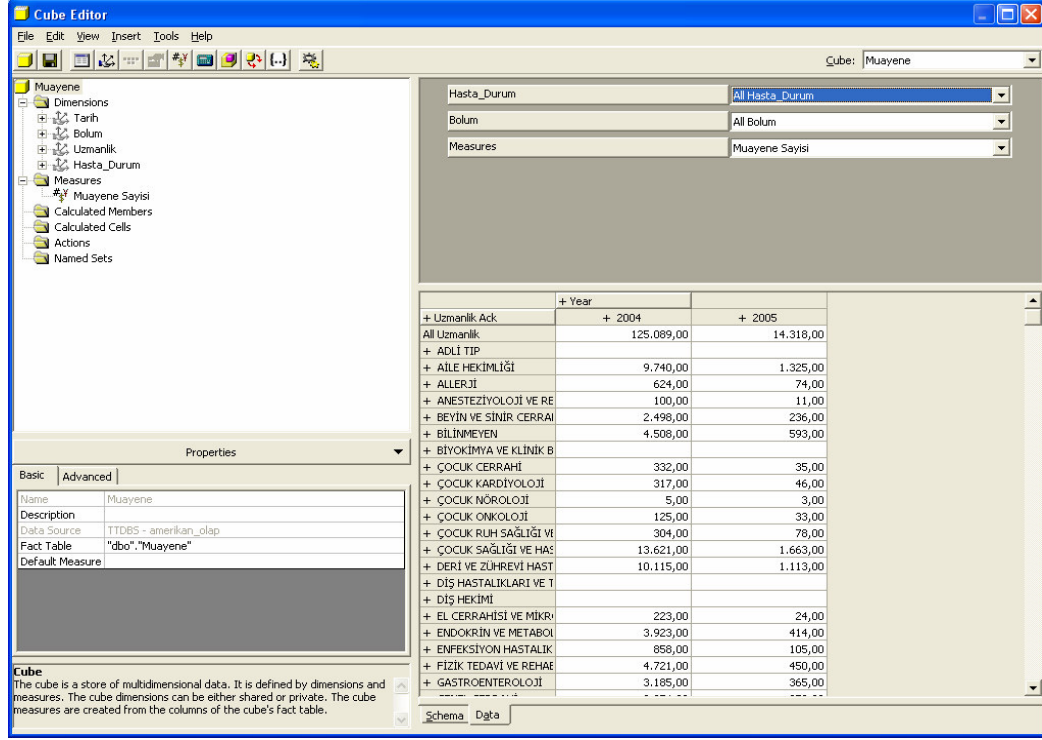
| <b>Küp Adı</b> | <b>Açıklama</b>  |
|----------------|--|
| Muayene        | Uzmanlık, doktor, bölüm, bölüm tipi, hasta durumu ve zaman boyutları ile muayene sayılarını gruplayan küp tanımıdır.     |
| Olum           | Ölüm nedenlerine ve zaman boyutuna göre ölen hasta sayılarını gruplayan küp tanımıdır.                                   |
| Patoloji       | Patoloji türleri, hasta durumu ve zaman boyutuna göre gerçekleşmiş patoloji sonuçlarının sayılarını veren küp tanımıdır. |
| Radyoloji      | Radyoloji işlem sayılarının tetkik türü, hasta durumu ve tarih boyutlarıyla ifade edildiği küp tanımıdır.                |
| Recete         | Zaman boyutuna göre verilmiş reçete sayılarının tutulduğu küp tanımıdır.   |
| Tıbbi_Islemler | Tıbbi işlem sayılarının işlem türü, hasta durumu, yapıldığı bölüm ve tarih boyutlarıyla ifade edildiği küp tanımıdır.    |
| Yatak          | Hastanede bulunan yatak sayılarının yatak türüne ve tarih boyutuna göre tutulduğu küp tanımıdır.                         |
| Yatışı_Yapılan | Uzmanlık dallarına ve tarih boyutuna göre yatışı yapılan hasta sayılarının tutulduğu küp tanımıdır.                      |

#### **4.2.5.3. Küp yapılarının özetlenmiş ve detay verileri tutacak şekilde işletilmesi**

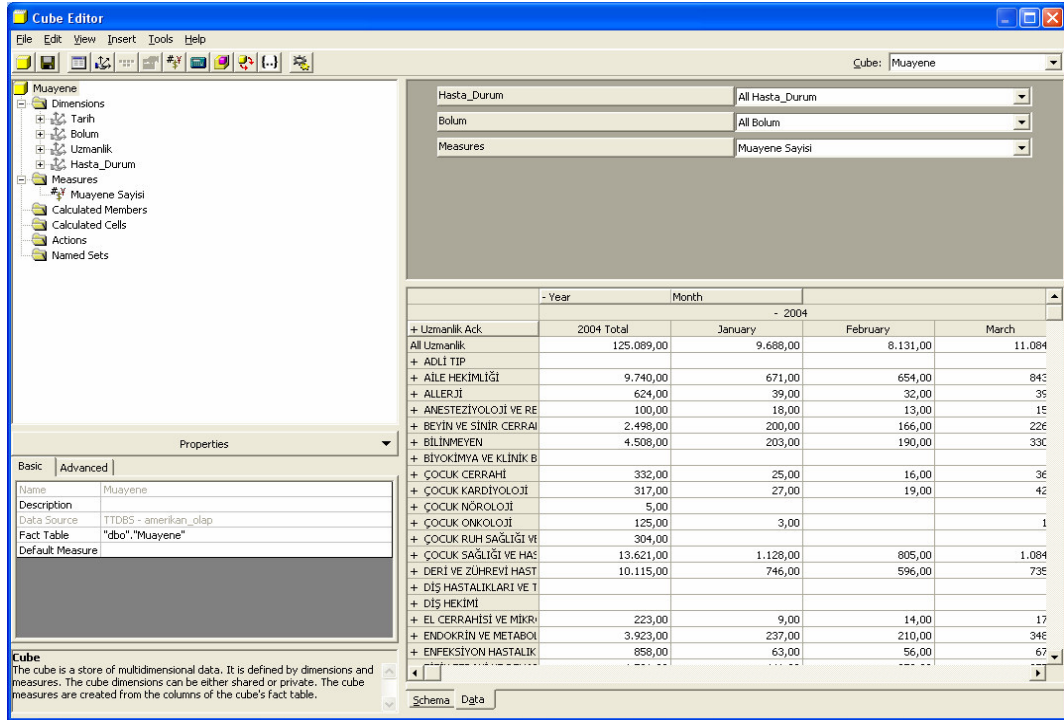
Tasarlanan küpler işletilerek MOLAP modeline göre veri ambarından çekilen veriler küp veri yapıları bünyesine aktarılmıştır. Veri ambarına eklenen her yeni veri için küplerin yeniden işletilerek güncel veriye sahip olması sağlanmalıdır. Bu kapsamda yapılan çalışmalar 4.2.6 nolu bölümde açıklanmıştır.

#### **4.2.5.4. Hazırlanan küpler üzerinde küp editörü üzerinden sorgulama yapılarak sistemin test edilmesi**

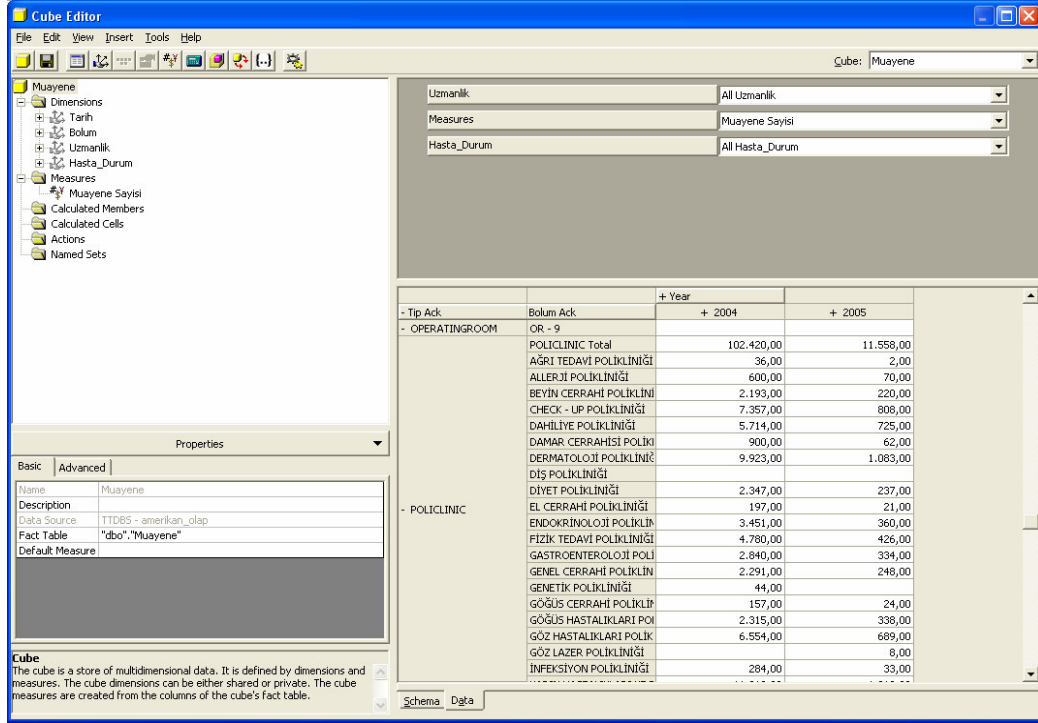
İşletilerek verilerle beslenen küpler üzerinde Microsoft Analysis Manager ürününde yer alan küp editörü ile küp içerikleri incelenebilir ve küpler üzerinde gerçekleşen “roll-up, drill-down, slice ve dice” işlemleri gerçekleştirilir. Şekil 4.12 – Şekil 4.16 kapsamında çeşitli küp editörü görünümleri yer almaktadır.



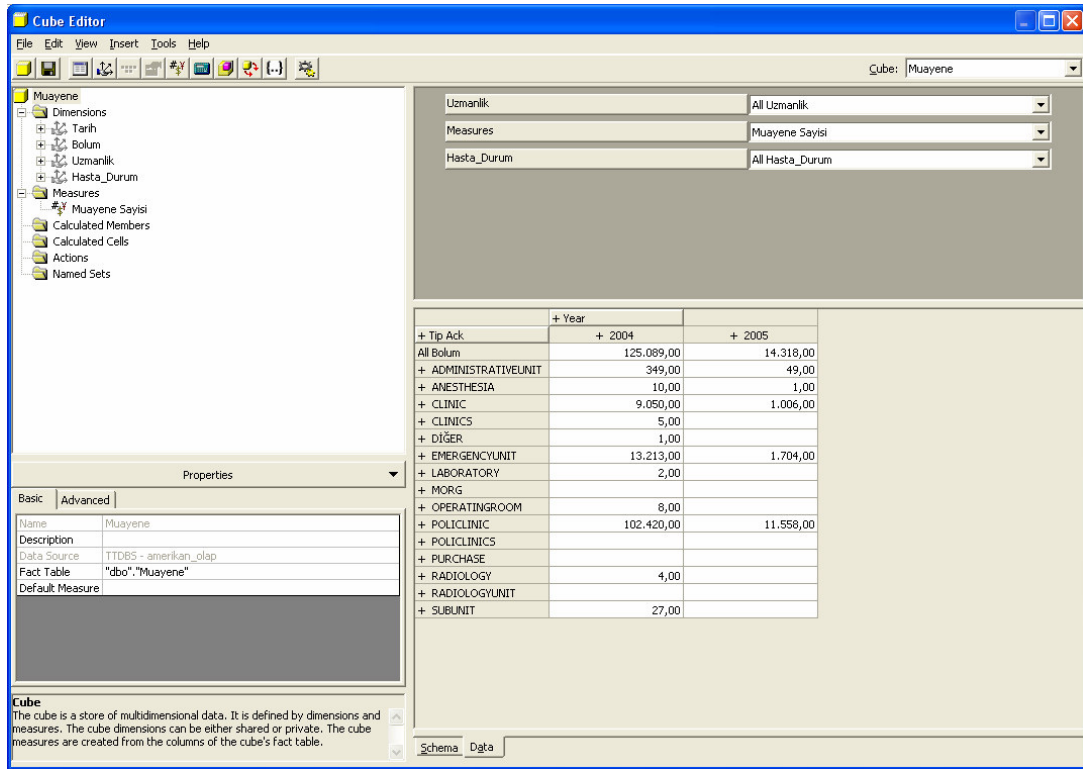
Şekil 4.12 Muayene Küpü kapsamında Uzmanlık Dalı satırı ve Tarih sütununa göre muayene sayılarını gösteren küp editörü görünümü



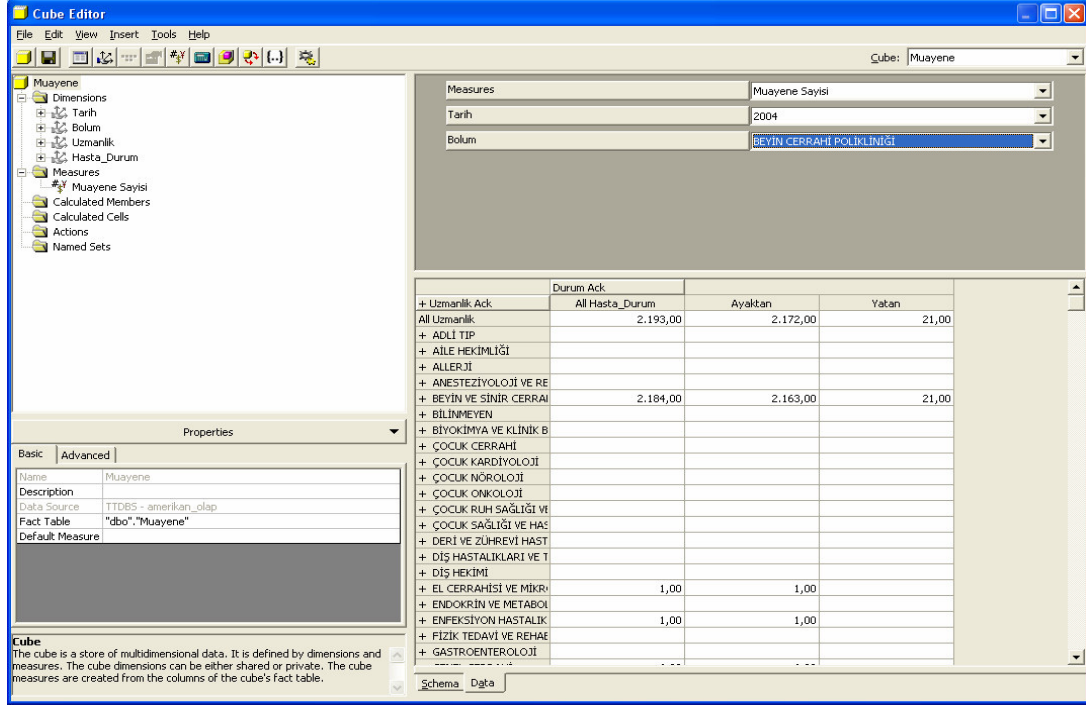
Şekil 4.13 Tarih boyutunun ayrıntı düzeyi artırılması (drill-down) sonucu küp editörü görünümü



Şekil 4.14 Muayene Küpü kapsamında Bölümler satırı ve Tarih sütununa göre muayene sayılarını gösteren küp editörü görünümü



Şekil 4.15 Muayene küpü üzerinde bölüm tiplerine göre gruplama seviyesi artırılması (roll-up) sonucu oluşan küp editörü görünümü

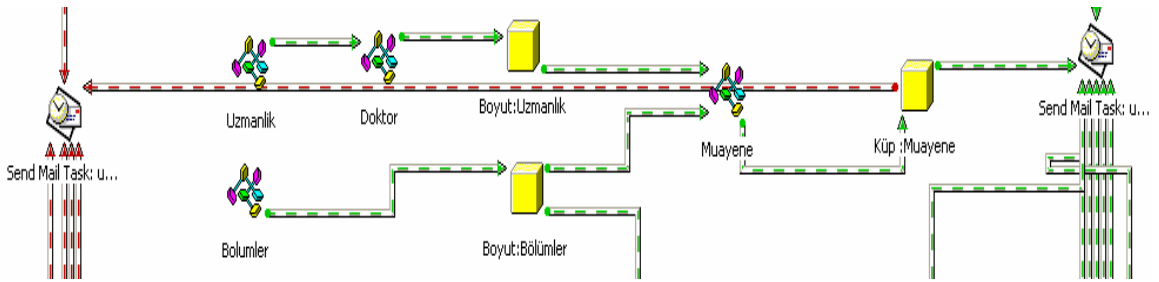


Şekil 4.16 2004 yılında, Beyin Cerrahisi Polikliniğinde yapılmış olan muayene sayılarının uzmanlık dalı ve hasta durumuna göre gruplanması sonucunu gösteren küp editörü görünümü (slice – dice)

#### 4.2.6. Veri ambarının güncelliğinin korunmasının sağlanması

Belirlenen bir zaman diliminde veri ambarında bulunan tabloların ÇİHİ sisteminde idame edilen güncel verilerle güncellenmesi sağlanmıştır. Öncelikle boyutların güncellenmesi gerekmektedir. Örneğin sisteme yeni bir bölüm eklenmesi durumunda öncelikle bu yeni kaydın veri ambarı üzerinde ilgili tabloya aktarılması ve sonrasında bu boyut tablosunu referans olarak oluşturulmuş boyutların güncellenmesi gerekir. Bu sayede küp boyutlarının güncel kalınması sağlanmış olur. Yeni eklenmiş olan bölüm kapsamında gerçekleşmiş olan tüm muayene bilgileri muayene küpü için birer olay bilgisidir. Öncelikle veri ambarında bulunan olay tablo ilgili veri ile güncellenir sonrasında çok boyutlu muayene küpünün işletilerek yeni verilere sahip olması sağlanır. Sistemin güncelliğinin korunması için yeni verilerin her gün 23:45 de düzenli olarak çalıştırılacak iş paketi bünyesinde bulunan betiklerin işletilmesiyle sağlanır. Şekil 4.10 ile bu kapsamda hazırladığım iş paketi bütününe ait görüntü yer almaktadır. Örnek olarak 'Muayene' küpünün güncellenmesini ele alalım, bu bölüme ait pakette yer alan görevler ve iş akışları

Şekil 4.17 ile gösterilmiştir. Muayene küpünü besleyen boyut tabloları, 'Uzmanlik, Bolum, Hasta\_Durum ve Tarih' tablolarıdır. 'Hasta\_Durum' tablosu 'Yatan ve Ayaktan' olmak üzere iki değeri taşıyan kod tablosu olduğu için güncelleme ihtiyacı bulunmamaktadır. 'Tarih' tablosu, her yıla ait tarih içerikleri tutan tablodur ve her yeni yıla ihtiyaç olduğunda ayrıca üretilmesi gerekmektedir. 'Uzmanlık' boyutu 'Doktor' ve 'Uzmanlik' tablolarından, 'Bolum' boyutu 'Bolumler' ve 'Bolum\_Tip' tablolarından üretilmiştir. Muayene küpünün güncellenmesinden önce uzmanlık ve bölüm boyutlarının güncellenmesine gereksinim duyulur. Bu nedenle Şekil 4.17 üzerinde görüleceği gibi, öncelikle uzmanlık kod tablosunun veri ambarında bulunmayan yeni verilerle güncellenmesini sağlayan betik işletilir. Bu işlem 'başarılı' olduğu zaman doktor tablosunun yeni verilerle güncellenmesini sağlayan betik devreye alınır. İşlemin başarı ile sonuçlanması durumunda Uzmanlık boyutu yeni verilere sahip olacak şekilde işletilir. Bölüm verileri için de benzer şekilde öncelikle veri kaynağından güncel veriler veri ambarına alınır, daha sonrasında Bolum boyutu güncellenir. Her iki boyut içinde tanımlanmış işlemlerin başarılı olması durumunda Muayene bilgilerine ait yeni işlemler veri kaynağından veri ambarına ilgili betiğin çalıştırılmasıyla aktarılır. Bu işlemin de başarılı olması durumunda Muayene küpü işletilerek güncel verilere sahip olması sağlanmış olur. Küpün işletilmesinden sonra sistem yöneticisine Muayene küpünün işletilmesinin başarıyla sonuçlandığını bildiren bir elektronik posta iletisinin gönderilmesi sağlanır. Herhangi bir aşamada oluşan başarısızlık durumunda da sistem yöneticisine başarısızlığın olduğu adımı içeren bir elektronik posta iletisinin gönderilmesi sağlanır.



Şekil 4.17 Veri Dönüştürüm İş Paketi Bünyesinde Bulunan ve Muayene Küpünün Güncellenmesini Sağlayan Görevlere Ait İş Akışı

### 4.3. Karar Ağacı Algoritması İle Veri Madenciliği Modelinin Tanımlanması

Microsoft Analysis Services ürününün sağladığı karar ağacı algoritması kullanılarak örnek veri madenciliği modelleri hazırlanmıştır. Hazırlanan modeller veri kaynağı olarak hem veri ambarını hem de ÇİAİ küplerini kullanmaktadır. Karar ağacı algoritması, ameliyat geçirmiş hasta bilgileri üzerinde uygulanmıştır. Tabloda yer alan toplam kayıt sayısı 2125. Tüm kayıtlar veri ambarına aktarılırken temizlenmiş ve düzenlenmiştir. Bu kapsamda yapılan adımlar;

- Nitelik alanlarının boş olmaması sağlanmıştır. Nitelik alanlarından herhangi birisi boş olan kayıtlar veri ambarına aktarılmamıştır.
- Doğum yılı niteliği için, doğum tarihi bilgi alanının yıl hanesinde bulunan içerik değeri aktarılmıştır.
- Cinsiyet niteliği için, 'E', 'Erkek', 'K', 'Kadın', 'Bay', 'Bayan' gibi çeşitli şekillerde tutulan bilgi içeriği 'E' ve 'K' değerlerini ifade edecek şekilde düzenlenmiştir.
- Meslek niteliği için aynı mesleği ifade etmek için kullanılan birden çok içerik bilgisi tek bir açıklama altında toplanmıştır. Örneğin Bilgisayar Mühendisi için, 'Bil. Müh, BILMUH, Bilg. Mhd., Bilgisayar Mühendisi, Bilgisayar Müh. Bilg. Mühendisi' gibi açıklamalar 'Bilgisayar Müh' olarak düzenlenmiştir.
- Hastanın geçirdiği ameliyat bilgisi, 'Ameliyat\_Tur' tablosunda yer alan ameliyat kodlarını içermektedir. Kod içeriği yerine ameliyatı ifade eden açıklama bilgisi aktarılmıştır. Ayrıca tabloda yer alan toplam kayıt sayısının azlığı nedeniyle en alt seviyede tutulan ameliyat içerikleri yerine bir üst seviyede yer alan ameliyat grup içerikleri aktarılmıştır. Bu sayede, aynı ameliyat grubuna giren hasta kümelerinin oluşması sağlanmıştır.

Verilerin tutulduğu ilişkisel tablo yapısı Çizelge 4.6 ve karar ağacı algoritmasında kural oluşturmak için kullanılacak nitelik değerlerinin içeriği Çizelge 4.7 ile verilmiştir. Veri madenciliği modelinden beklediğim sonuç ameliyat türlerine göre ilgili ameliyatların gerçekleşme yüzdelerinin cinsiyet, eğitim durumu, medeni hal, doğum yılı, hasta grubu ve statü bilgilerine göre gruplanmasıydı. Fakat elimde bulunan kayıtların sayıca çok yeterli olmaması ve ameliyat türlerin çok geniş bir spekturuma sahip olmasından ötürü beklediğim sonuçlara ulaşamadım.



Kullandığım veriler belirli bir ameliyat grubuna ait olsaydı, karar destek kapsamında çok daha anlamlı sonuçlar izleyebilirdim.

Çizelge 4.6 Hasta Bilgileri Tablo Yapısı

| <b>Nitelik Adı</b> | <b>İçerik Türü</b> |
|--------------------|--------------------|
| Hasta No           | Sayısal Alan       |
| Statü              | Metin Alanı        |
| Cinsiyet           | Metin Alanı        |
| Medeni Hali        | Metin Alanı        |
| Doğum Yılı         | Metin Alanı        |
| Meslek             | Metin Alanı        |
| Eğitim Durumu      | Metin Alanı        |
| Hasta Grubu        | Metin Alanı        |
| Ameliyat Açıklama  | Metin Alanı        |

Çizelge 4.7 Hasta Bilgileri İçin Kural Oluşturulabilecek Nitelikler

| <b>Nitelik</b> | <b>Olası Değerler</b>   |
|----------------|---|
| Statü          | VIP, Normal, Personel, Felaket, Takip, Yabancı  |
| Cinsiyet       | E, K  |
| Medeni Hali    | Dul, Evli, Bekar  |
| Eğitim Durumu  | İlköğretim, Ortaöğretim, Lise, Yüksekokul, Üniversite, Lisansüstü, Okur Yazar Değil, Okur Yazar             |
| Doğum Yılı     | Sürekli   |
| Hasta Grubu    | Firma Hastası, Olağanüstü Durum Hastası, Kamu Kurum Hastası, Personel, Bireysel Hasta, Özel Sigortalı Hasta |

Çizelge 4.7 Hasta Bilgileri İçin Kural Oluşturulabilecek Nitelikler (devam)

| Nitelik               | Olası Değerler   |
|-----------------------|--|
| Meslek                | Araştırma Grv., Avukat, Ayakkabıcı, Bankacı, Basın, Bilgisayar Müh., Danışman, Biyolog, Emekli, Elektrik Müh., Eczacı, Doktor, Ev Hanımı, Fizik Müh., Fırıncı, Grafiker, İç Mimar, Memur, Matbaacı, Sigortacı, Yönetici ...<br>(Toplam 171 Meslek Açıklaması bulunmaktadır.) |
| Ameliyat Açıklama     | 511 adet ameliyat açıklaması bulunmaktadır.  |
| Üst Ameliyat Açıklama | 511 adet ameliyat açıklamasının bağlı olduğu üst ameliyat grubunu ifade eder. 113 adet ameliyat grubu bulunmaktadır.   |

#### 4.3.1. İlişkisel Veri Ambarı Üzerinde Karar Ağacı Algoritması

İlişkisel veri ambarında bulunan hasta bilgilerinin tutulduğu 'Hasta' tablosu kullanılarak veri madenciliği modeli hazırlanmıştır. Bu modelin hazırlanmasında girdi ve tahmin edilebilir amaçlı kullanılan nitelikler Çizelge 4.8 ile verilmiştir. Bir niteliğin 'Girdi' olarak kullanılması, karar ağacı algoritmasında yer alacak kuralların belirlenmesinde girdi verisi olması anlamını taşımaktadır. Bir niteliğin 'Tahmin Edilebilir' olarak kullanılması, ilgili nitelik için karar ağacının tanımlanması ve bu niteliğin belirlenecek kurallara göre gerçekleşme oranlarının tespit edilmesi anlamını taşımaktadır. Her 'Tahmin Edilebilir' nitelik için bir karar ağacı oluşturulacaktır. Çizelge 4.8 üzerinde görüldüğü gibi 'Cinsiyet' ve 'Üst Ameliyat Açıklama' nitelikleri için birer karar ağacı hazırlanacaktır.

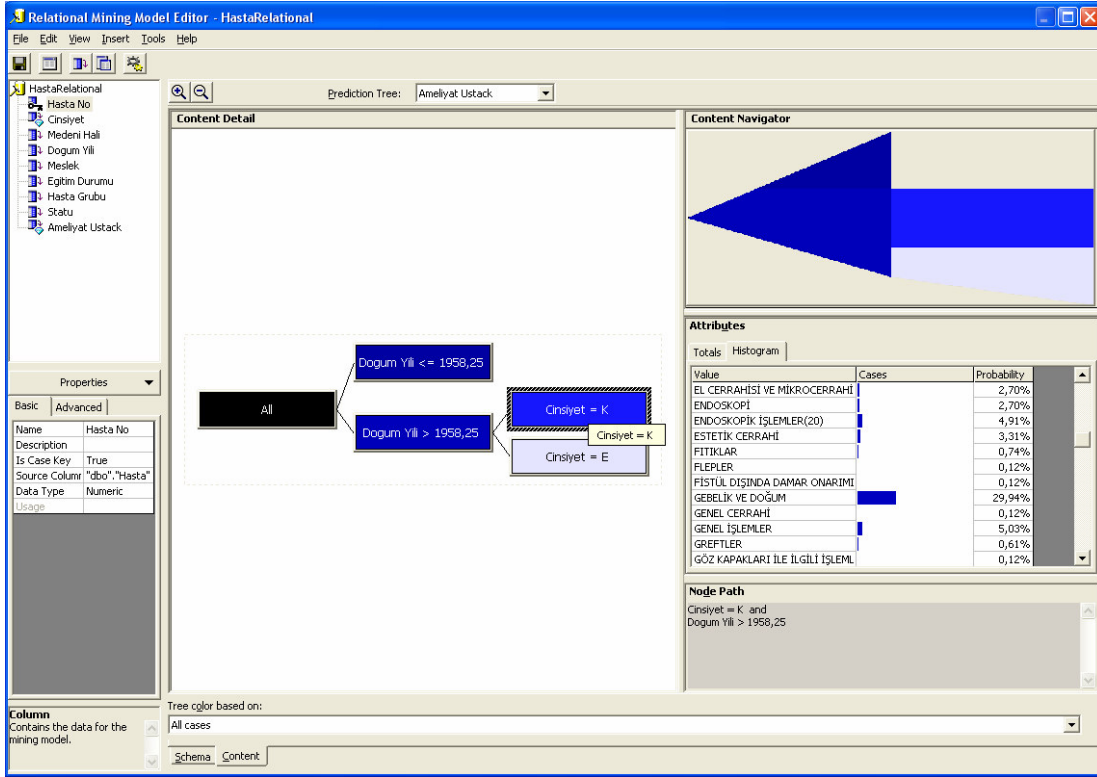
Çizelge 4.8 Veri Madenciliği Modelinde Kullanılan Nitelikler ve Türleri

| Nitelik               | Türü (Girdi, Tahmin Edilebilir) |
|-----------------------|---------------------------------|
| Cinsiyet              | Girdi, Tahmin Edilebilir        |
| Üst Ameliyat Açıklama | Girdi, Tahmin Edilebilir        |
| Medeni Hali           | Girdi                           |
| Doğum Yılı            | Girdi                           |
| Meslek                | Girdi                           |
| Eğitim Durumu         | Girdi                           |
| Hasta Grubu           | Girdi                           |
| Statü                 | Girdi                           |

Şekil 4.18 ile gösterildiği gibi 'Hasta' ilişkisel tablosunda bulunan 'Üst Ameliyat Açıklama' niteliği için üretilen karar ağacı iki basamaklıdır. İlk basamakta 'Doğum Yılı' niteliğinin 1958,25 ten küçük ve eşit olduğu durum ile 'Doğum Yılı' niteliğinin 1958,25 ten büyük olduğu durum ele alınmıştır. İkinci seviye ise 'Cinsiyet' niteliğinin 'K' ve 'E' değerlerine göre oluşturulmuştur. Elde edilen karar ağacındaki kurallar Çizelge 4.9 ile verilmiştir.

Çizelge 4.9 Üst Ameliyat Açıklama' niteliği üretilen karar ağacı kuralları

| Kurallar                               | Örnek Sonuç Kümesi (Ameliyat Grupları)  |
|--|---|
| Doğum Yılı $\leq$ 1958,25              | El cerrahisi ve Mikro Cerrahi , %3,5,<br>Genel İşlemler %7,49, Fıtıklar %4,8,<br>Karaciğer ve Safra Yolları %3,58                               |
| Doğum Yılı $>$ 1958,25 ve Cinsiyet = K | El cerrahisi ve Mikro Cerrahi , %2,70,<br>Genel İşlemler %5,03, Fıtıklar %0,74,<br>Karaciğer ve Safra Yolları %2,21,<br>Gebelik ve Doğum %29,94 |
| Doğum Yılı $>$ 1958,25 ve Cinsiyet = E | El cerrahisi ve Mikro Cerrahi , %3,07,<br>Genel İşlemler %7,55, Fıtıklar %1,42,<br>Karaciğer ve Safra Yolları %1,42                             |

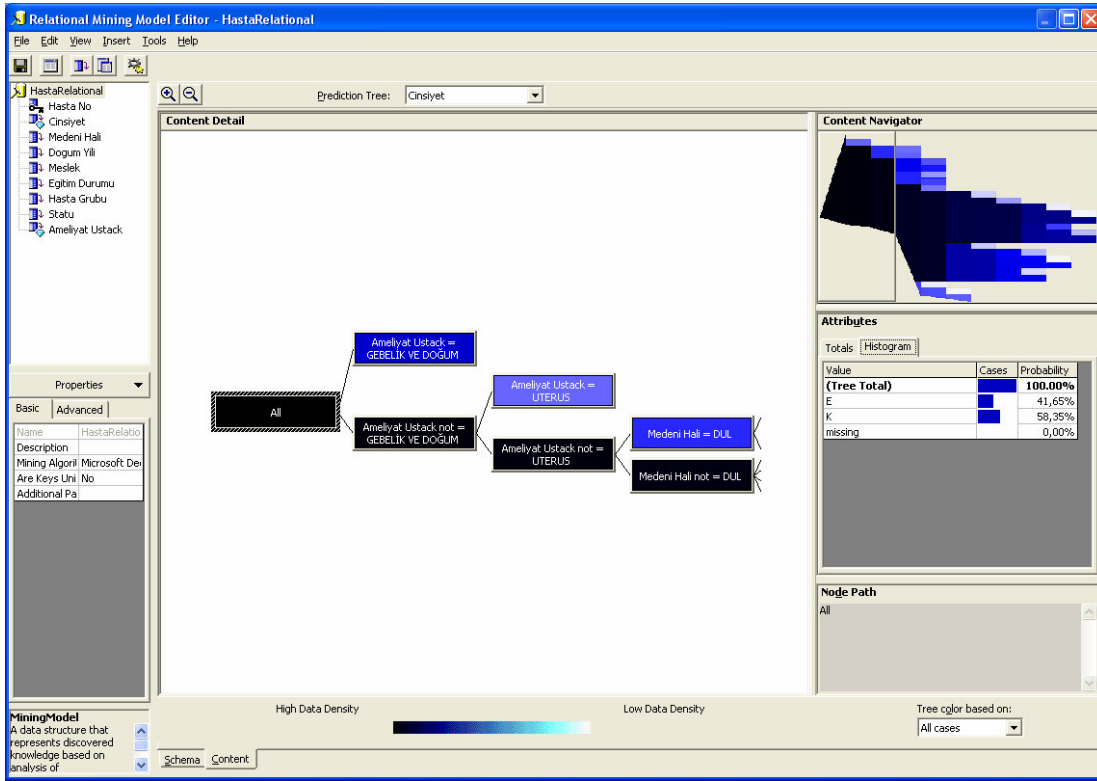


Şekil 4.18 İlişkisel Veri Ambarı Tablosu İle 'Üst Ameliyat Açıklama' Niteliği İçin Oluşturulan Karar Ağacı

Şekil 4.19 ile gösterildiği gibi 'Hasta' ilişkisel tablosunda bulunan 'Cinsiyet' niteliği için üretilen karar ağacı bir çok seviyeden oluşmaktadır. İlk basamakta 'Üst Ameliyat Açıklama' niteliğinin 'Gebelik ve Doğum' içeriğine sahip olup olmama durumu ele alınmıştır. Elde edilen karar ağacındaki bazı kurallar Çizelge 4.10 ile verilmiştir.

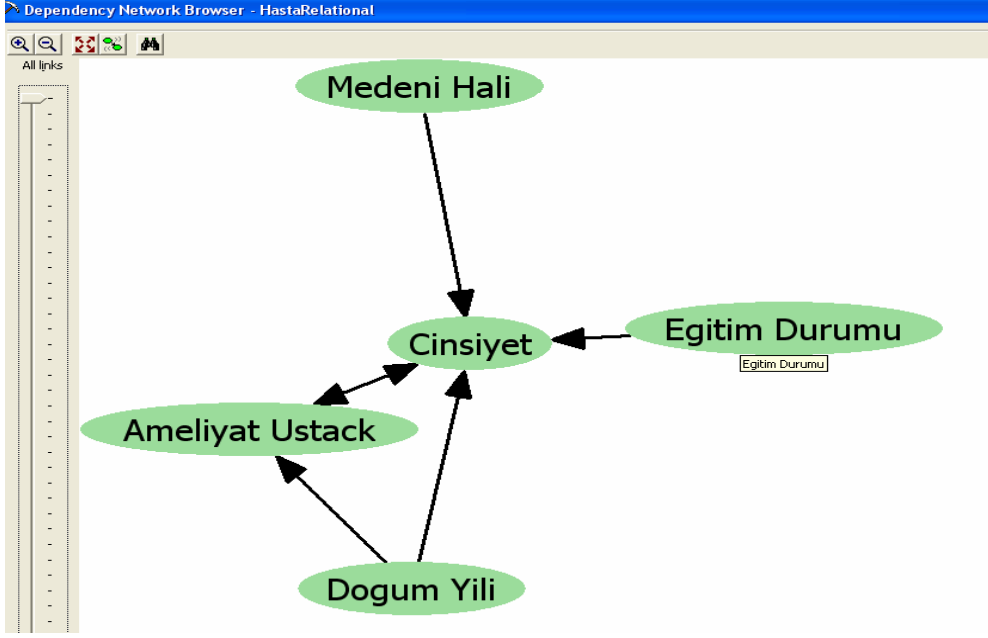
Çizelge 4.10 'Cinsiyet' niteliği üretilen karar ağacı kuralları örnekleri

| <b>Kurallar</b>   | <b>Örnek Sonuç Kümesi (Cinsiyet)</b> |
|---|--------------------------------------|
| Üst Ameliyat = Gebelik ve Doğum   | Erkek %0,40, Kadın %99,19            |
| Üst Ameliyat not = Gebelik ve Doğum ve Üst Ameliyat = Uterus  | Erkek %2,04, Kadın %95,92            |
| Medeni Hali = Dul ve Ameliyat Ustack not = Gebelik ve Doğum ve not = Uterus   | Erkek %50, Kadın %49,94              |
| Medeni Hali not = Dul ve Ameliyat Ustack not = Gebelik ve Doğum ve not = Uterus   | Erkek %86,73, Kadın %12,24           |
| Medeni Hali = Dul ve Eğitim Durumu = Lise ve Ameliyat Ustack not = Gebelik ve Doğum ve not = Uterus   | Erkek %2,22, Kadın %95,56            |
| Medeni Hali not = Dul ve Doğum Yılı > 1932,25 ve <= 1935,25 ve Ameliyat Ustack not = Gebelik ve Doğum ve not = Uterus   | Erkek %86,73, Kadın %12,24           |
| Medeni Hali not = Dul ve Doğum Yılı > 1967,25 ve <= 1987,25 ve Ameliyat Ustack not = Gebelik ve Doğum ve not = Uterus ve not = Endoskopik İşlemler ve not = Estetik Cerrahi   | Erkek %34,26, Kadın %65,46           |
| Medeni Hali not = Dul ve Doğum Yılı > 1935,25 ve <= 1967,25 ve Eğitim Durumu not = Üniversite ve Ameliyat Ustack not = Gebelik ve Doğum ve not = Uterus ve not = Endoskopik İşlemler ve not = Prostat ve not = Meme | Erkek %42,58, Kadın %57,21           |



Şekil 4.19 İlişkisel Veri Ambarı Tablosu İle 'Cinsiyet' Niteliği İçin Oluşturulan Karar Ağacı

Karar ağacında oluşturmak için kullanılan niteliklerin birbirine bağımlılık durumunu belirten 'Bağımlılık Ağı – Dependency Network' çizeneği Şekil 4.20 ile gösterilmiştir. Bu çizenekten de anlaşılacağı gibi 'Cinsiyet' niteliği 'Medeni Hal, Eğitim Durumu, Doğum Yılı ve Ameliyat Üst Açıklama' niteliklerine bağımlıdır. 'Ameliyat Üst Açıklama' niteliği de 'Cinsiyet ve Doğum Yılı' niteliklerine bağımlıdır. 'Hasta Grubu, Statü ve Meslek' nitelikleri girdi niteliği olarak belirtilmiş olsa da verilerin dağılımından dolayı bu niteliklerin karar ağacı oluşmasında bir etkisi bulunmamaktadır.



Şekil 4.20 Bağımlılık Ağı Çizeneği

#### 4.3.2. ÇİAİ Küpleri Üzerinde Karar Ağacı Algoritması

İlişkisel veri ambarında bulunan hasta bilgileri ÇİAİ kapsamında çok boyutlu küp veri yapıları ile ifade edilecek şekilde ayarlanmıştır. Oluşturulan ÇİAİ küpü kullanılarak karar ağacı algoritması uygulanarak veri madenciliği modeli oluşturulmuştur.

##### 4.3.2.1. Hasta Bilgileri ÇİAİ Küpü

İlişkisel veri ambarında bulunan hasta bilgileri kayıtları ÇİAİ kapsamında 'HastaBlg' isimli küp veri yapısına aktarılmıştır. HastaBlg küpüne ait ölçüm değeri 'Ameliyat Sayısı' dır. Küpün boyutları Çizelge 4.11 ile verilmiştir.

Çizelge 4.11 HastaBlg Kúpünde Yeralan Boyut Tanımları

| Boyut Adı     | İşlevi  |
|---------------|---|
| Hasta_Numara  | İlişkisel veri tablosunda bulunan her bir kayıt temel anahtar olarak 'Hasta_No' adlı bir sütun ile ifade edilmektedir. 'Hasta_Numara' boyutu temel anahtar konumundaki bu alan üzerinden tanımlanmış yıldız şema boyutudur. Bu boyut üzerinde, ilgili hasta kaydı kapsamında cinsiyet, medeni hal, doğum yılı, meslek, eğitim durumu, statü, hasta grubu ve ameliyat açıklaması nitelikleri için üye tanımları yapılmıştır. |
| Ameliyat_Tur  | 'Ameliyat' kúpü için de kullanılan ve ameliyat türlerinin baba-oğul sıradüzensel yapısına sahip boyut tanımıdır.  |
| Cinsiyet      | Hasta bilgileri tablosunda yer alan kayıtlar arasında bulunan cinsiyet bilgilerine göre tanımlanmış boyut tanımıdır.  |
| Medeni_Hal    | Hasta bilgileri tablosunda yer alan kayıtlar arasında bulunan medeni hal bilgilerine göre tanımlanmış boyut tanımıdır.  |
| Eğitim_Durumu | Hasta bilgileri tablosunda yer alan kayıtlar arasında bulunan eğitim durumu bilgilerine göre tanımlanmış boyut tanımıdır.   |
| Meslek        | Hasta bilgileri tablosunda yer alan kayıtlar arasında bulunan meslek bilgilerine göre tanımlanmış boyut tanımıdır.  |

HastaBlg kúpüne ait kúp yapısı ve kúp editörü üzerindeki veri içerikleri Şekil 4.21 ile gösterilmiştir. Cinsiyet, Medeni\_Hal, Eğitim\_Durumu ve Meslek boyutları, tanımlanan kúp üzerinde bu kriterlerle sorgulama yapılabilmesini sağlamak için tanımlanmıştır. Bu boyutlar kendi boyut içerikleri kapsamında toplam ameliyat sayılarını sağladıkları için veri madenciliği modelinde kullanılmamıştır. Veri madenciliği modeli kapsamında Hasta\_Numara boyutunda tutulan üye tanımları kullanılmıştır. Her bir üye tanımı, Hasta\_Numara boyutunda yer alan her bir kayıt için değer tuttuğundan dolayı madencilik modeline uygun detayda veri sağlayabilmektedir.



The screenshot shows the 'Cube Editor' interface for a cube named 'HastaBlg'. The left pane displays a hierarchical tree of dimensions: 'HastaBlg' (Dimensions) -> 'Hasta\_Numara' (Dimensions) -> 'Hasta No' (Dimensions) -> 'Member properties' (Dimensions) -> 'Cinsiyet' (Dimensions), 'Medeni Hal' (Dimensions), 'Dogum Yili' (Dimensions), 'Statu' (Dimensions), 'Meslek' (Dimensions), 'Egitim Durumu' (Dimensions), 'Hasta Grubu' (Dimensions), 'Ameliyat Ustack' (Dimensions). Other dimensions include 'Ameliyat\_Tur' (Dimensions), 'Cinsiyet' (Dimensions), 'Medeni Hal' (Dimensions), 'Egitim Durumu' (Dimensions), 'Meslek' (Dimensions). The 'Measures' section includes 'Ameliyat Sayisi' (Measures), 'Calculated Members' (Measures), 'Calculated Cells' (Measures), 'Actions' (Measures), and 'Named Sets' (Measures). The right pane shows a pivot table with the following data:

|  | MeasuresLevel | Cinsiyet        |       |
|--|---------------|-----------------|-------|
|  |               | Ameliyat Sayisi |       |
|  | All Cinsiyet  | E               | K     |
| + En Ust   |               |                 |       |
| Tüm Ameliyatlara                                   | 2.125         | 885             | 1.240 |
| + 3-ALGOLOJİ-AĞRI TEDAVİLERİ                       |               |                 |       |
| + AMELİYATLAR                                      | 140           | 64              | 76    |
| + ARTRODEZ   |               |                 |       |
| + ÇOCUK CERRAHİSİ(5)                               | 41            | 24              | 17    |
| + DAMAR CERRAHİSİ                                  | 55            | 41              | 14    |
| + GENEL CERRAHİ                                    | 439           | 217             | 222   |
| + GÖĞÜS CERRAHİSİ                                  | 17            | 14              | 3     |
| + GÖZ HASTALIKLARI                                 | 48            | 27              | 21    |
| + KADIN HASTALIKLARI VE DOĞUM                      | 392           | 3               | 389   |
| + KALP VE DAMAR CERRAHİSİ                          | 37            | 21              | 16    |
| + KULAK-BURUN-BOĞAZ HASTALIKLARI                   | 177           | 94              | 83    |
| + KURUMSAL PAZARLAMA                               |               |                 |       |
| + LAZER TEDAVİ UYGULAMALARI                        | 3             | 2               | 1     |
| + LENF NODLARI VE LENF DÜĞÜMLERİ                   |               |                 |       |
| + NÖROŞİRURJİ                                      | 117           | 57              | 60    |
| + ORTOPEDİ VE TRAVMA                               | 277           | 131             | 146   |
| + ORTOPEDİK ONKOLOJİ                               | 31            | 15              | 16    |
| + PLASTİK, REKONSTRÜKTİF VE YAKINLAŞTIRICI CERRAHİ | 177           | 60              | 117   |
| + REKONSTRÜKSİYONLARI                              | 1             | 1               |       |

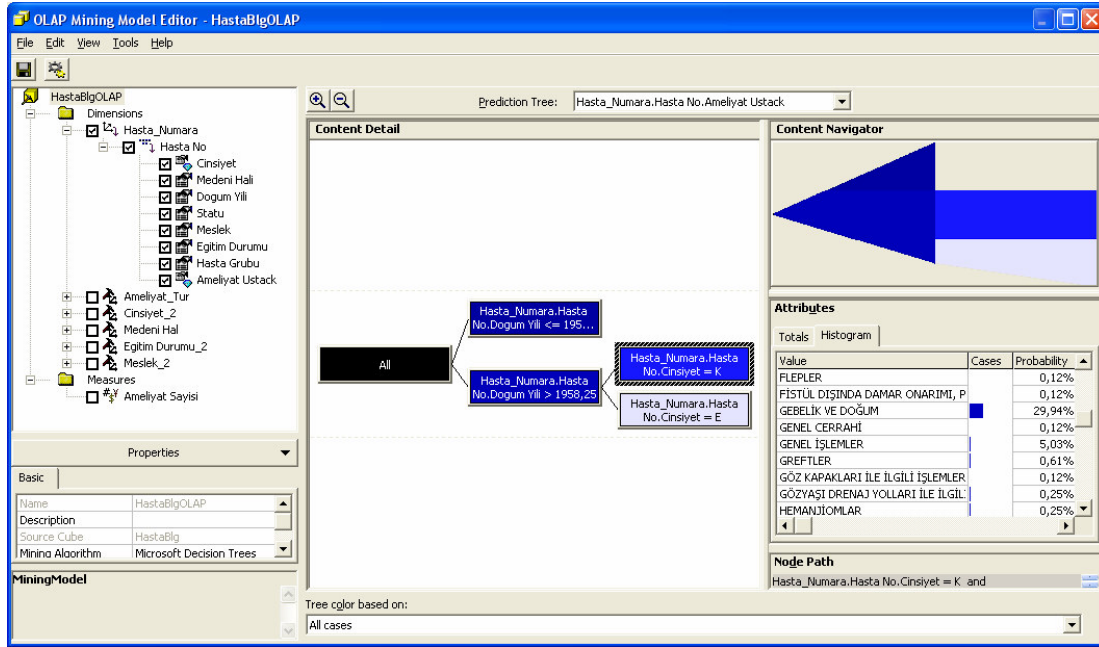
Şekil 4.21 HastaBlg Küp İçeriği

#### 4.3.2.2. Hasta Bilgileri ÇİAİ Küpü ile Karar Ağacı Algoritmasının Uygulanması

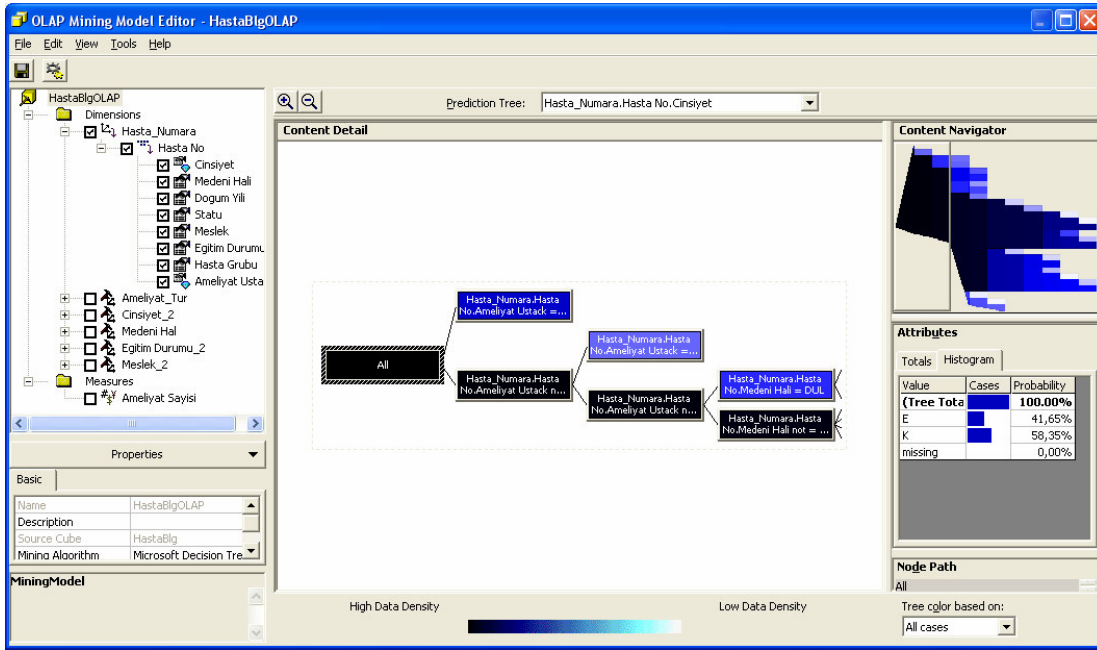
HastaBlg küpü kapsamında veri madenciliği modeli için kullanılan boyut ve üye tanımları Çizelge 4.12 ile gösterilmiştir. 'Tahmin Edilebilir' nitelik olarak seçilen 'Cinsiyet ve Üst Ameliyat Açıklama' üye tanımları için ayrı ayrı karar ağaçları oluşturulmuştur. Elde edilen karar ağaçları Şekil 4.22 ve Şekil 4.23 ile gösterilmiştir. Veri kaynağının ilişkisel veri ambarı ya da ÇİAİ küpü olarak belirlenmesinde elde edilecek karar ağacının kuralları ve görünümü kapsamında hiçbir farklılık olmadığı gözlemlenmiştir. ÇİAİ küpünün veri kaynağı olarak seçilmesi durumunda elde edilen karar ağacı sanal küp olarak saklanabilmektedir ve ağacı oluşturan her bir kural bir boyut üyesi gibi düşünülerek sorgulanabilmektedir. Oluşturulan sanal küpe ait görüntü Şekil 4.24 ile verilmiştir.

Çizelge 4.12 Veri Madenciliği Modelinde Kullanılan Boyutlar ve Üye Tanımları

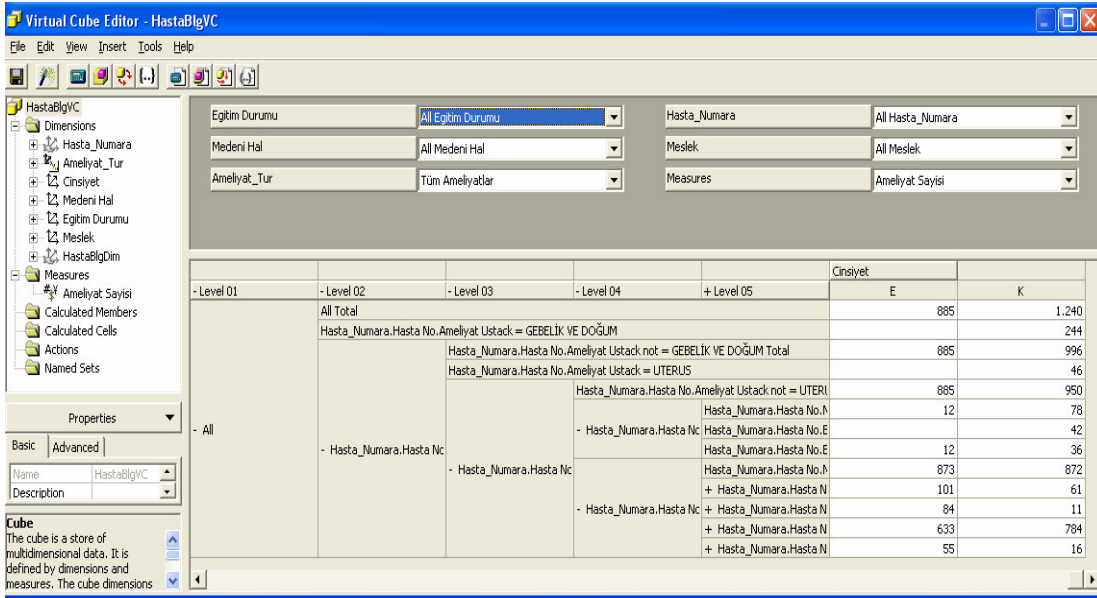
| Boyut        | Üye Tanımı            | Türü (Girdi, Tahmin Edilebilir) |
|--------------|-----------------------|---------------------------------|
| Hasta_Numara | Cinsiyet              | Girdi, Tahmin Edilebilir        |
| Hasta_Numara | Üst Ameliyat Açıklama | Girdi, Tahmin Edilebilir        |
| Hasta_Numara | Medeni Hali           | Girdi                           |
| Hasta_Numara | Doğum Yılı            | Girdi                           |
| Hasta_Numara | Meslek                | Girdi                           |
| Hasta_Numara | Eğitim Durumu         | Girdi                           |
| Hasta_Numara | Hasta Grubu           | Girdi                           |
| Hasta_Numara | Statü                 | Girdi                           |



Şekil 4.22 'Ust Ameliyat Açıklama' Üye Tanımı İle Oluşturulan Karar Ağacı



Şekil 4.23 'Cinsiyet' Üye Tanımı İle Oluşturulan Karar Ağacı



Şekil 4.24 Karar Ağacını Gösteren Sanal Küp İçeriği

#### 4.4. Karar Destek Amaçlı Sorgulama ve Raporlama Aracı Gerçekleştirimi

Sorgulama ve raporlama aracının çalışabileceği örnek bir veri ambarı, ÇİAİ küpleri ve karar ağacı algoritması ile tanımlanmış veri madenciliği modeli oluşturulduktan sonra raporlama aracının gerçekleştirilmesine geçilmiştir. Microsoft Visual Studio .NET 2003 geliştirme ortamı kullanılarak yazılım üretilmiştir. Bu yazılım aracı ile

karar destek amacıyla karar alıcı konumundaki kişiler için, ÇİAİ küpleri ve veri madenciliği modelleri ile tutulan veri kümelerinin anlaşılır arayüzler içerisinde sunulması ve bir ya da daha fazla sorgu sonucunu bünyesinde barındırabilen karar destek raporlarının üretilmesi sağlanmıştır. Sorgulama dili olarak ÇBİ cümlecikleri kullanılmıştır. ÇİAİ küplerinin sorgulanabilmesi için uygun ÇBİ cümleciklerinin üretilmesi sağlanmıştır. Hazırladığım yazılım ürünü bünyesinde yer alan ana başlıklar aşağıda yer almaktadır.

#### **4.4.1. Veri Ambarı ve ÇİAİ Küplerine Erişim Bağlantısının Sağlanması**

Veri ambarı ilişkisel veri tabanı olarak Microsoft SQL Server VTYS bünyesinde yer almaktadır. ÇİAİ küpleri Microsoft Analysis Services bünyesinde hazırlanmıştır. İki farklı ortamda tutulan verilere erişmek için iki ayrı erişimi sağlayacak bağlantı parametrelerine ihtiyaç duyulmaktadır. Bu parametreler veri kaynağının bulunduğu sunucunun adresi ya da ismi, bağlantı sağlayıcı türü, bağlantı için gerekli kullanıcı adı ve şifresi, erişilecek katalog ismidir. ÇİAİ sunucusuna bağlanmak için kullanılan bağlantı cümlecigi şu şekildedir;

Data Source = localhost; Provider = MSOLAP; integrated security = SSPI; Initial Catalog = OLAPKaynak

Veri kaynağı olarak belirtilen "localhost" ifadesi, yazılımın çalıştığı yerel bilgisayar ortamını ifade etmektedir. Yani ÇİAİ sunucusu kullanılan yerel makine üzerinde çalışmaktadır. MSOLAP bağlantı sağlayıcıyı ifade eder. "integrated security = SSPI" ifadesi bağlantı için işletim sistemine bağlantıda kullanılan kullanıcı adı ve şifresinin geçerli olacağını belirtir. Aksi takdirde ,UId = "Kullanıcı\_adi"; Pwd = "Kullanıcı\_Şifre" şeklinde parametreler kullanılabilirdi. Bağlantı cümleciginde yer alan en son parametre ile de, ÇİAİ küplerinin bulunduğu kataloğun isminin "OLAPKaynak" olduğu belirtilmiştir.

Veri ambarının yer aldığı ilişkisel veritabanı sunucusuna bağlanmak için kullanılan bağlantı cümlecigi şu şekildedir;

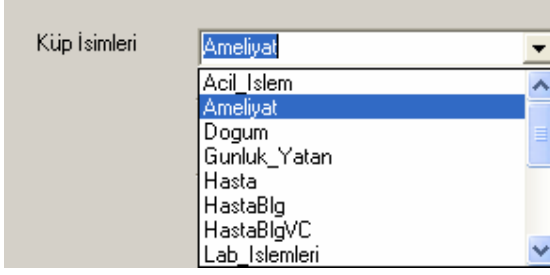
Data Source = localhost; Provider = SQLOLEDB.1; integrated security = SSPI ; Initial Catalog= VAKaynak

Bağlantı cümleciğinde bağlantı sağlayıcı SQLOLEDB.1 olarak değişmiştir. Bu değişiklikle bir veri tabanı bağlantı sağlayıcısı seçilmiş olur. “VAKaynak” olarak adlandırılmış ilişkisel veritabanına, işletim sistemine bağlantıda kullanılan kullanıcı adı ve şifresi kullanılarak erişilmektedir.

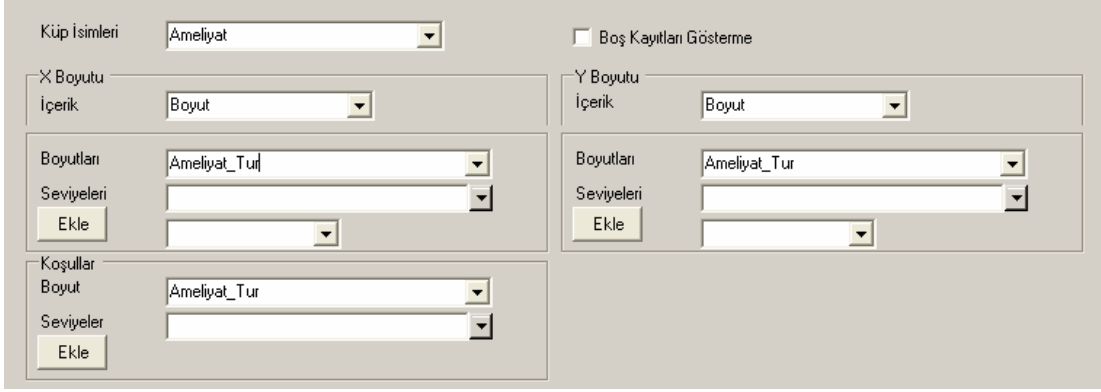
Veri ambarı ve ÇİAİ sunucularına erişim için kullanılan bağlantı cümleciklerinde yer alan parametreler değiştirilerek fiziksel olarak farklı bilgisayarlarda bulunan sunuculara ve farklı kataloğlara erişim sağlanabilmektedir. Bağlantı cümlecikleri, hazırlanmış olan yazılım içerisinde gömülmemiş, ‘app.config’ şeklinde adlandırılmış bir uygulama konfigürasyon dosyasında tutulmuştur. Bu sayede, yazılım üzerinde hiç bir değişiklik yapmadan veri ambarı ve ÇİAİ sunucularında yapılabilecek fiziksel adres değişiklikleri ya da kullanılan kataloğlarda oluşacak değişiklikler kolaylıkla ele alınabilmektedir. Tez kapsamında kullandığım örnek veri ambarı ve ÇİAİ küpleri ‘Hastane Bilgi Yönetim Sistemi’ verilerini içerse de, hazırladığım yazılım her çeşit veri kataloğu üzerinde çalışabilmektedir.

#### **4.4.2. Küp Seçimi**

Karar destek amaçlı sorgulama ve raporlama aracı çalıştırıldığında ilk olarak uygulama konfigürasyon dosyası okunarak gerekli bağlantı cümleciklerine erişip sunuculara bağlanılmaktadır. ÇİAİ sunucusundan hizmet olarak belirtilen katalog altında yer alan küpler, küplere ait boyutlar ve her boyutta bulunan seviyeler araç üzerinde gösterilmektedir. Sorgulama için ÇBİ cümleciği üretilerek sonuç kümesi grid bileşeninde gösterilmektedir. Öncelikle sorgulama yapılacak küpün seçilmesi gerekmektedir. Sistem çalışmaya başladığında, ÇİAİ sunucusuna erişerek mevcut küplerin bir liste veri yapısı üzerinde saklanmasını sağlamaktadır. Küp tanımları alfabetik sıraya göre liste üzerinde tutulmaktadır. Sorgulanmak istenilen küp bu liste üzerinden seçilerek, seçilen küpe ait boyut bilgilerinin ÇİAİ sunucusundan çekilmesi sağlanmaktadır. Küp tanımlarının yer aldığı liste veri yapısı Şekil 4.25 ile gösterilmiştir. Şekil 4.26 üzerinde de görüldüğü gibi “Boyutları” başlığı ile ifade edilmiş liste yapıları seçilmiş olan küpe ait boyutları göstermektedir. “Seviyeleri” başlığı ile ifade edilmiş liste yapıları seçilmiş olan boyuta ait seviyeleri içermektedir.



Şekil 4.25 Sistemde Yeralan Küplerin Tutulduğu Liste Yapısı



Şekil 4.26 Sistemde Yeralan Küplerin, Boyutların ve Seviyelerin Tutulduğu Liste Yapıları

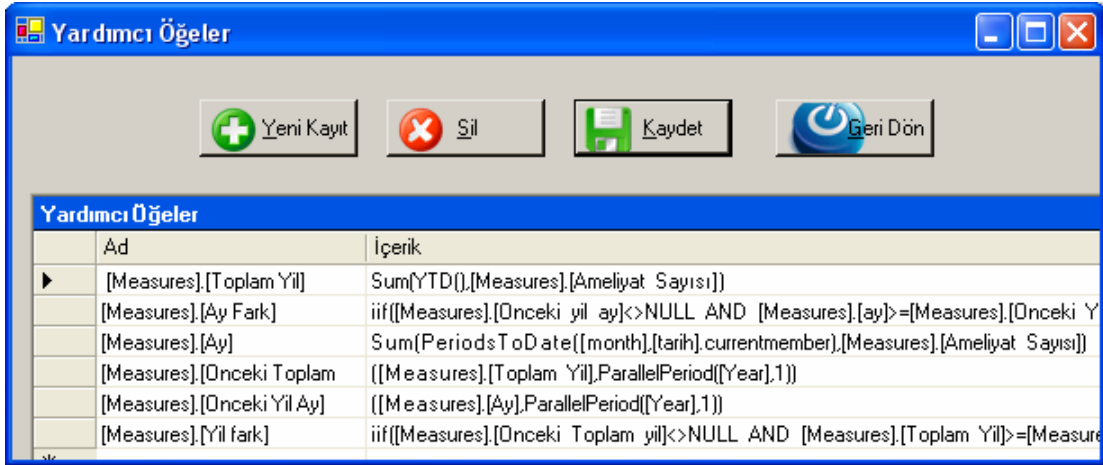
### 4.4.3. ÇBİ Cümlecığının Üretilmesi

ÇBİ Cümlecığı 2.9 nolu bölümde de açıklandığı gibi belirli bir küp üzerinde sütun ve satır olarak boyutların belirtildiği cümleciklerdir. ÇBİ cümlecığının oluşturulabilmesi için öncelikle bir küp seçilmeli, daha sonra satır ve sütunları ifade edecek boyut kümeleri belirtilmelidir. Satır ve sütunlarda bir boyut seviyesi ya da bir boyut üyesi bulunabileceği gibi tanımlanmış hesaplama öğeleri de kullanılabilir. Hesaplama öğeleri yardımcı öğeler olarak adlandırılmıştır.

#### 4.4.3.1. Boyut – Seviye ya da Yardımcı Öğe Seçimi

ÇBİ cümlecikleri satır ve sütun değerlerinin seçilmesi ile çok boyutlu bir küpün iki boyuta indirgenerek incelenmesini sağlayan sorgular üretir. Dolayısıyla ÇBİ cümleciklerinde hem satır hem de sütun değeri için ayrı ayrı boyut ya da hesaplanmış öğeler belirlenmelidir. Her bir sütun içeriği grid bileşeni üzerinde X boyutu olarak, her bir satır içeriği de Y boyutu olarak gösterilir. Hazırlanan

uygulama yazılımı kapsamında hem X hem de Y boyutu için her bir boyutta görüntülenmesi istenilen boyut ya da hesaplanmış öğeler seçilmelidir. Şekil 4.26 üzerinde de görüldüğü gibi 'X Boyutu ve Y Boyutu' olarak adlandırılmış bölümler bulunmaktadır. 'İçerik' başlığı altında 'Boyut ya da Yardımcı Öğeler' olmak üzere iki seçenek bulunmaktadır. 'Boyut' seçeneğinin seçilmesi durumunda ilgili küp için mevcut boyutların ve boyut seviyelerinin seçilmesi sağlanır. Hesaplanan öğeler 2.9.1.6 nolu başlık altında açıklanmıştır. 'Yardımcı Öğeler' seçeneğinin seçilmesi durumunda bir ilişkisel veri tablosu üzerinde saklanan hesaplanmış öğelerden seçim yapılması sağlanır. Yardımcı Öğeler, 'Diğer İşlemler' menüsünün altında yer alan 'Yardımcı Öğeler' başlığıyla erişilen veri giriş ve güncelleme formu kapsamında sisteme kaydedilir. Yardımcı öğeyi tanımlayan 'Ad' ve içerdiği ÇBİ cümlecini belirten 'İçerik' alanlarından oluşan kayıtlar tutulur. Bu formun yapısı Şekil 4.27 ile gösterilmiştir.



Şekil 4.27 Yardımcı Öğeler Veri Giriş ve Güncelleme Ekranı

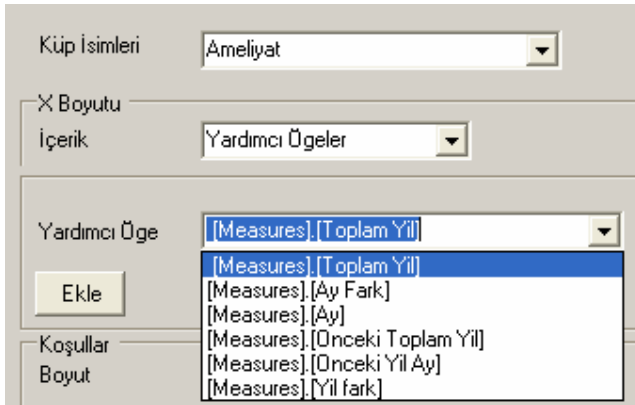
İçerik olarak 'Boyut' bilgisi alındığında 'Boyutları' ve 'Seviyeleri' başlıklı iki liste oluşmaktadır. 'Boyutları' başlıklı liste kapsamında, seçilmiş olan küpte yer alan boyutlar listelenir. Örnek olarak 'Ameliyat' küpü seçildiğinde bu küpte kullanılan 'Ameliyat Tur, Anestezi, Hasta Durum ve Tarih' boyutları listelenecektir. Bu listeden 'Ameliyat Tur' boyutu seçildiğinde 'Seviyeleri' başlıklı liste ameliyat türlerine göre hazırlanmış boyutta yer alan seviyeleri ağaç yapısında görüntüleyecek şekilde güncellenir. Bu durum Şekil 4.28 ile gösterilmiştir. Seçilen seviyeyi belirten ÇBİ komutu liste üzerinde oluşturulmaktadır. Göz hastalıkları

bünyesindeki 'Anjiografi' seviyesi için '[Ameliyat\_Tur].[Tüm Ameliyatlar].[GÖZ HASTALIKLARI].[ANJIOGRAFI]' içerikli ÇBİ komutu elde edilir.



Şekil 4.28 Ameliyat Tür Boyutu İçin Boyut Seviyeleri Listesi

İçerik olarak 'Yardımcı Öğeler' bilgisi alındığında 'Yardımcı Öğe' başlıklı liste ekranda görüntülenmektedir. Bu liste kapsamında, önceden veri tabanına kaydedilmiş hesaplanmış öğelere ait tanım başlıkları bulunmaktadır. Seçilen öğe başlığına göre içerik bilgisine karşılık gelen ÇBİ komutu, ÇBİ cümlecğine eklenecektir. Örneğin, '[Measures].[Toplam Yıl]' yardımcı öğesinin belirlenmesiyle, 'Sum(YTD(),[Measures].[Ameliyat Sayısı])' hesaplanmış öğe içeriği seçilmiş olacaktır. Seçilen hesaplanmış öğe içeriği YTD – (Year-To-Date) ve Sum işlevlerini içeren ve kısıt olarak verilecek tarih içeriğini yıl bilgisine dönüştürerek elde edilen yıl için toplam ameliyat sayıları hesaplamaktadır.



Şekil 4.29 Yardımcı Öğe Listesi



#### 4.4.3.2. İşlem Seçimi

İşlem seçimi boyut kullanımı için geçerlidir. Desteklenen işlem türleri;

- **Hiçbir işlem seçilmeme durumu;** Seçilmiş olan boyut için ilgili boyuttaki tüm seviyelerin ya da seçilmiş seviyenin ÇBİ cümleciğine eklenmesini sağlar. Sadece boyutun seçilmesi durumunda 'members' komutu ile ilgili boyutun tüm seviyelerinin sorguya eklenmesi sağlanır. Eğer bir seviye seçilmişse, seçilen seviyenin ÇBİ komut içeriği ÇBİ cümleciğine eklenir. Örneğin, ameliyat türlerine ait tüm seviyeler için '[Ameliyat\_Tur].members' komutu, 'Genel Cerrahi' altında yer alan 'Genel İşlemler' seviyesi için '[Ameliyat\_Tur].[Tüm Ameliyatlar].[GENEL CERRAHİ].[GENEL İŞLEMLER]' komutu eklenecektir.
- **Ve (,);** belirlenmiş olan boyut üzerindeki bazı seviyelerin seçilerek satır ya da sütun olarak ifade edilmesini sağlar. Örneğin, tarih boyutu kapsamında 2004 yılı için Ocak, Şubat ve Mayıs aylarının satır ya da sütun değeri olarak kullanılması istendiğinde her bir seviye tek tek yazılarak virgül ile birleştirilir; [Tarih].[2004].[Ocak], [Tarih].[2004].[Şubat], [Tarih].[2004].[Mayıs]
- **Arasında (:)** ; belirlenmiş olan boyut kapsamındaki bir seviye aralığının seçilmesi sağlanır. Örneğin tarih boyutu için 2004 yılı Ocak ve Haziran ayları arasındaki tüm aylar istendiğinde  
[Tarih].[2004].[Ocak] : [Tarih].[2004].[Haziran]  
şeklinde ':' karakteri ile istenilen seviye aralığı belirtilir.
- **Birleşme (Crossjoin);** birden çok boyut içeriğinin satır ya da sütun olarak ifade edilmesini sağlar. Tek bir birleşme işlemi ile sadece iki adet boyutun birleştirimi sağlanır. Daha fazla boyut kullanımı gerekiyorsa, iç içe yazılacak birleşme komutları ile ÇBİ cümleciği türetilir. Satır ya da sütun üzerinde birden çok boyuta ait verilerin kartezyen çarpım mantığı ile gösterilmesi sağlanır. Örneğin;

```
'select { Crossjoin( [LabTetkik].members ,[Hasta_Durum].members) } on  
columns, { [Tarihler].[2004].[Ocak] : [Tarihler].[2004].[Temmuz] } on rows
```

from Lab\_Islemleri' ifadesi ile sütun kapsamında laboratuvar tetkik ve hasta durum bilgilerinin yer alması istenmektedir. Her bir laboratuvar tetkiki için hasta durum bilgileri tekrarlanmaktadır. Bu ifadeye ait sonuç kümesinin grid bileşeni ile ifadesi Şekil 4.30 ile gösterilmiştir.

|         | Alanin Amino Transferaz (ALT) |          |        | Albumin |         |        | Albumin, biyolojik sıvı |         |       |
|---------|-------------------------------|----------|--------|---------|---------|--------|-------------------------|---------|-------|
|         | Tüm                           | Ayaktan  | Yatan  | Tüm     | Ayaktan | Yatan  | Tüm                     | Ayaktan | Yatan |
| Ocak    | 1.527,00                      | 976,00   | 551,00 | 254,00  | 133,00  | 121,00 | 2,00                    |         | 2,00  |
| Şubat   | 1.348,00                      | 889,00   | 459,00 | 269,00  | 141,00  | 128,00 | 1,00                    |         | 1,00  |
| Mart    | 1.904,00                      | 1.328,00 | 576,00 | 407,00  | 251,00  | 156,00 |                         |         |       |
| Nisan   | 1.864,00                      | 1.266,00 | 598,00 | 440,00  | 265,00  | 175,00 |                         |         |       |
| Mayıs   | 1.923,00                      | 1.231,00 | 692,00 | 442,00  | 239,00  | 203,00 | 1,00                    |         | 1,00  |
| Haziran | 1.851,00                      | 1.234,00 | 617,00 | 428,00  | 238,00  | 190,00 |                         |         |       |
| Temmuz  | 1.876,00                      | 1.284,00 | 592,00 | 446,00  | 265,00  | 181,00 | 1,00                    | 1,00    |       |

Şekil 4.30 Sütun Kapsamında Birleştirme İşlemi Örneği

- **Sıradüzensel Gösterim;** baba-oğul ilişkisine sahip seviyelendirilmiş boyutlar için seviyelerin sıradüzensel yapıda gösterimi sağlanır. Bu sayede verinin anlaşılabilirliği artırılmış olur. Alt seviyelere sahip olan seviyeler '+' işareti ile ifade edilir. Örneğin, Ameliyat küpü üzerinde ameliyat türleri ve tarih boyutları ile sorgulama yapmak istediğimizde; sütun kapsamında tüm ameliyat türlerini ve satır kapsamında tarihlerin yer almasını sağlayan ÇBİ cümlecisi

```
'select { [Ameliyat_Tur].members } on columns, { [Tarihler].members } on rows from Ameliyat'
```

şeklinde olacaktır. Bu ifadenin çalıştırılması, ameliyat türlerinin çok fazla olmasından dolayı oldukça zaman almaktadır ve elde edilen sonuç kümesinin anlaşılabilirliği oldukça azdır. Bu ifadeyi sıradüzensel yapıda oluşturmak ve grid bileşeninin özelliklerinden yararlanarak alt seviyelerin ihtiyaç duyulduğunda gösterimini sağlamak çok daha kullanışlı bir yöntem olmaktadır. Aynı zamanda 'drill-down' ve 'roll-up' işlemlerinin yapılmasını sağlayarak anlaşılabilirliği artırmaktadır. Sıradüzensel olarak hazırlanan ÇBİ cümlecisi

```
'select {[Ameliyat_Tur].levels(0).members, Ameliyat_Tur].levels(1).members } on columns, { [Tarihler].members } on rows from Ameliyat'
```

şeklindedir. Bu sorgu ile ameliyat türleri kapsamında ilk seviyede bulunan ameliyat türlerine ait sonuçlar kullanıcıya sunulmaktadır. Kullanıcı, ihtiyaç duyduğu detayda bilgiye ulaşmak için grid bileşeni üzerinde fare ile çift tıklayarak istenilen seviyenin bir alt seviyesine inebilmektedir.

|         | EPIDURAL HASTA KONTROLLU ANALJEZİ (EPIDURAL) | EPIDURAL KAN/SERUM YAMASI (AMELİYAT) | EPIDUROSKOPİ (AMELİYAT) | FASET EKLEM (TEK) RFT (AMELİYAT) | FASET EKLEM İCİ (LÖMBER) ENJEKSİYON (AMELİYAT) | FASET EKLEM İCİ (SERVİKAL) ENJEKSİYON (AMELİYAT) | FASET MEDIAN SINIR LUMBAR (TEK) (AMELİYAT) |
|---------|--|--------------------------------------|-------------------------|----------------------------------|--|--|--|
| 2004    |  | 2,00                                 |                         | 1,00                             | 10,00  |  |  |
| Ocak    |  |                                      |                         |                                  | 2,00   |  |  |
| Şubat   |  |                                      |                         | 1,00                             | 3,00   |  |  |
| Mart    |  | 1,00                                 |                         |                                  |  |  |  |
| Nisan   |  |                                      |                         |                                  | 1,00   |  |  |
| Mayıs   |  |                                      |                         |                                  | 2,00   |  |  |
| Haziran |  |                                      |                         |                                  |  |  |  |
| Temmuz  |  |                                      |                         |                                  |  |  |  |
| Ağustos |  | 1,00                                 |                         |                                  |  |  |  |
| Eylül   |  |                                      |                         |                                  | 2,00   |  |  |

Şekil 4.31 Sıradüzensel gösterim kullanılmadan elde edilen sonuç kümesi

|         | Tüm Ameliyathar | 3-ALGOLOJİ+AĞRI TEDAVİSİ UYGULAMALARI | + AMELİYATLAR | + ARTRODEZ | + ÇOCUK CERRAHİSİ(5) | + DAMAR CERRAHİSİ | + GENEL CERRAHI | + GÖĞÜS CERRAHİSİ | + GÖZ HASTALIKLARI | + KADIN HASTALIKLARI VE DOĞUM |
|---------|-----------------|---------------------------------------|---------------|------------|----------------------|-------------------|-----------------|-------------------|--------------------|-------------------------------|
| 2004    | 10.329,00       |                                       | 485,00        | 4,00       | 334,00               | 212,00            | 1.785,00        | 88,00             | 241,00             | 1.943,00                      |
| Ocak    | 862,00          |                                       | 41,00         |            | 33,00                | 12,00             | 148,00          | 10,00             | 15,00              | 184,00                        |
| Şubat   | 733,00          |                                       | 33,00         | 1,00       | 28,00                | 15,00             | 114,00          | 2,00              | 12,00              | 147,00                        |
| Mart    | 1.148,00        |                                       | 51,00         |            | 28,00                | 50,00             | 204,00          | 14,00             | 20,00              | 188,00                        |
| Nisan   | 1.056,00        |                                       | 42,00         | 2,00       | 27,00                | 18,00             | 192,00          | 8,00              | 34,00              | 207,00                        |
| Mayıs   | 945,00          |                                       | 52,00         |            | 18,00                | 21,00             | 164,00          | 5,00              | 24,00              | 189,00                        |
| Haziran | 827,00          |                                       | 38,00         |            | 52,00                | 22,00             | 141,00          | 3,00              | 26,00              | 141,00                        |
| Temmuz  | 647,00          |                                       | 36,00         |            | 28,00                | 5,00              | 78,00           | 2,00              | 10,00              | 139,00                        |
| Ağustos | 730,00          |                                       | 27,00         |            | 39,00                | 8,00              | 130,00          | 5,00              | 20,00              | 146,00                        |
| Eylül   | 840,00          |                                       | 53,00         |            | 28,00                | 18,00             | 144,00          | 10,00             | 17,00              | 147,00                        |

Şekil 4.32 Sıradüzensel gösterim ile elde edilen sonuç kümesi

|       | Tüm Ameliyathar | 3-ALGOLOJİ+AĞRI TEDAVİSİ UYGULAMALARI | + AMELİYATLAR | - ARTRODEZ |   |  |   | + ÇOCUK CERRAHİSİ(5)                          | + DAMAR CERRAHİSİ  |
|-------|-----------------|---------------------------------------|---------------|------------|---|--|---|---|--------------------|
|       |                 | 3 ALGOLOJİ AĞRI TEDAVİSİ UYGULAMALAR  | AMELİYATLAR   | ARTRODEZ   | + ANTERİOR VEYA TEROLATERAL YAKLAŞIM TEKNİĞİ (SERVİKAL) | + ANTERİOR VEYA EROLATERAL YAKLAŞIM TEKNİĞİ (KAL-LÖMBER) | + POSTERİOR TEROLATERAL VEYA LATERAL TRANSVERS PROSES TEKNİĞİ | + POSTERİOR EROLATERAL VEYA LATERAL TRANSVERS | ÇOCUK CERRAHİSİ(5) |
| 2004  | 10.329,00       |                                       | 485,00        | 4,00       | 1,00  |  | 3,00  | 334,00  | 212,00             |
| Ocak  | 862,00          |                                       | 41,00         |            |   |  |   | 33,00   | 12,00              |
| Şubat | 733,00          |                                       | 33,00         | 1,00       |   |  | 1,00  | 28,00   | 15,00              |
| Mart  | 1.148,00        |                                       | 51,00         |            |   |  |   | 28,00   | 50,00              |

Şekil 4.33 Sıradüzensel gösterimle elde edilen sonuç kümesi üzerinde bir alt seviyeye erişim ile oluşan sonuç kümesi

- **Satır Birleşimi(UnionRow) ve Sütun Birleşimi (UnionColumn)**

Birden çok sorgunun satır ya da sütun bazında birleşik gösterimini sağlayan ve standart ÇBİ sözdizim yapısında yer almayan iki komut tez kapsamında tanımlanmış ve gerçekleştirilmiştir. 'UnionRow' ve 'UnionColumn' anahtar kelimeleri kullanılarak sorguların birleştirimi sağlanmıştır. Bu işlemlere ait detaylar aşağıdaki bölümlerde verilmiştir.

- **Sütun Birleşimi (UnionColumn)**

Sütun başlığı aynı olmak üzere iki sorgunun birleştirilerek tek sorgu gibi yorumlanarak tek bir sonuç kümesi üretmesi sağlanır. İstenilen sayıda sütun birleştirme işlemi yapılabilir. Grid üzerinde sunulacak sorgu sonuçlarının 'X' eksenini aynı boyut(lar) ile ifade edilerek birçok sorgu sonucunun alt alta üretilmesi sağlanmaktadır. Örnek olarak 'Laboratuar İşlemleri (Lab\_Islemleri)' küpü üzerinde iki farklı sorguyu ele alalım. Birinci sorgu, laboratuar tetkiklerinden 'Serbest Testesteron ve Toplam Testesteron' işlemlerini sütun bazında ve hasta durumlarını satır bazında oluştursun. Sorguya ait ÇBİ cümlecisi;

```
'Select { [LabTetiklik].[Tüm LabTetiklikler].[Testosteron, serbest]
,[LabTetiklik].[Tüm LabTetiklikler].[Testosteron, total] } on columns, {
[Hasta_Durum].members } on rows From Lab_Islemleri'
```

şeklinindedir. Bu sorguya ait sonuç kümesi Şekil 4.34 ile gösterilmiştir. İkinci sorgu, laboratuar tetkiklerinden 'Serbest Testesteron ve Toplam Testesteron' işlemlerini sütun bazında ve 2004 yılına ait tarih bilgilerini satır bazında oluştursun. Sorguya ait ÇBİ cümlecisi;

```
'Select { [LabTetiklik].[Tüm LabTetiklikler].[Testosteron, serbest],[LabTetiklik].[Tüm
LabTetiklikler].[Testosteron, total] } on columns, { [Tarihler].[2004].[Ocak] :
[Tarihler].[2004].[Aralık] } on rows From Lab_Islemleri'
```

olacaktır. Çalıştırılan ikinci ÇBİ cümlecisinin sonuç kümesi Şekil 4.35 ile verilmiştir.

|         | Testosteron, serbest | Testosteron, total |
|---------|----------------------|--------------------|
| Tüm     | 329,00               | 340,00             |
| Ayaktan | 325,00               | 331,00             |
| Yatan   | 4,00                 | 9,00               |

Şekil 4.34 Birinci Sorguya Ait Sonuç Kümesi

|         | Testosteron, serbest | Testosteron, total |
|---------|----------------------|--------------------|
| Ocak    | 9,00                 | 17,00              |
| Şubat   | 19,00                | 9,00               |
| Mart    | 48,00                | 32,00              |
| Nisan   | 37,00                | 30,00              |
| Mayıs   | 22,00                | 25,00              |
| Haziran | 38,00                | 24,00              |
| Temmuz  | 27,00                | 44,00              |
| Ağustos | 20,00                | 27,00              |
| Eylül   | 23,00                | 32,00              |
| Ekim    | 29,00                | 38,00              |
| Kasım   | 23,00                | 26,00              |
| Aralık  | 34,00                | 36,00              |

Şekil 4.35 İkinci Sorguya Ait Sonuç Kümesi

Her iki sorgunun sütun bileşeni aynıdır. Dolayısıyla iki sorguyu ayrı ayrı çalıştırıp sonuç kümesi üretmek yerine bu sorguları “Sütun Birleşimi (unioncolumn)” komutu ile birleştirerek çalıştırmak mümkündür. Yeni oluşacak ÇBİ cümleciği;

```
'Select { [LabTetik].[Tüm LabTetikler].[Testosteron, serbest]
,[LabTetik].[Tüm LabTetikler].[Testosteron, total] } on columns, {
[Hasta_Durum].members } on rows From Lab_Islemleri
```

#### **unioncolumn**

```
Select { [LabTetik].[Tüm LabTetikler].[Testosteron, serbest],[LabTetik].[Tüm
LabTetikler].[Testosteron, total] } on columns, { [Tarihler].[2004].[Ocak] :
[Tarihler].[2004].[Aralık] } on rows From Lab_Islemleri'
```

şeklinde olacaktır.

Yeni ÇBİ cümleciğinin çalıştırılması ile elde edilecek sonuç kümesine ait görüntü Şekil 4.36 ile gösterilmiştir. Öncelikle ilk sorgu çalıştırılarak grid

bileşenine sonuçlar aktarılmıştır. Daha sonra ikinci sorgu çalıştırılmıştır. İki sorguya ait sonuç kümesi tektir.

|           | Testosteron, serbest | Testosteron, total |
|-----------|----------------------|--------------------|
| Tüm Hasta | 329,00               | 340,00             |
| Ayaktan   | 325,00               | 331,00             |
| Yatan     | 4,00                 | 9,00               |
| Ocak      | 9,00                 | 17,00              |
| Şubat     | 19,00                | 9,00               |
| Mart      | 48,00                | 32,00              |
| Nisan     | 37,00                | 30,00              |
| Mayıs     | 22,00                | 25,00              |
| Haziran   | 38,00                | 24,00              |
| Temmuz    | 27,00                | 44,00              |
| Ağustos   | 20,00                | 27,00              |
| Eylül     | 23,00                | 32,00              |
| Ekim      | 29,00                | 38,00              |
| Kasım     | 23,00                | 26,00              |
| Ortalık   | 34,00                | 36,00              |

Şekil 4.36 Sütun Birleşimi Yapılmış Sorguya Ait Sonuç Kümesi

- **Satır Birleşimi (UnionRow)**

Satır içerikleri aynı olmak kaydıyla birden çok sorgunun yan yana çıktı üretecek şekilde çalışmasını sağlayan komuttur. İstenilen sayıda satır birleşimi yapılabilmektedir. Sırayla her bir sorgu çalıştırılarak grid bileşeni üzerinde yan yana sonuç üretmesi sağlanır. Sütun birleşimi kapsamında verilen örneği satır birleşimi ile oluşturalım. ‘Laboratuar İşlemleri(Lab\_Islemleri)’ küpü üzerinde iki tane sorgu oluşturalım. Birinci sorgu, laboratuar tetkiklerinden ‘Serbest Testesteron’ işlemini sütun bazında ve hasta durumlarını satır bazında oluştursun. Sorguya ait ÇBİ cümlecisi;

```
‘Select { [LabTetikik].[Tüm LabTetikikler].[Testosteron, serbest] } on columns, { [Hasta_Durum].members } on rows From Lab_Islemleri’
```

şeklindedir. Bu sorguya ait sonuç kümesi Şekil 4.37 ile gösterilmiştir. İkinci sorgu, laboratuar tetkiklerinden ‘Toplam Testesteron’ işlemini sütun bazında ve hasta durumlarını satır bazında oluştursun. Sorguya ait ÇBİ cümlecisi;

```
‘Select { [LabTetikik].[Tüm LabTetikikler].[Testosteron, total] } on columns, { [Hasta_Durum].members } on rows From Lab_Islemleri’
```

şeklindedir. Bu sorguya ait sonuç kümesi Şekil 4.38 ile gösterilmiştir. Her iki sorguyu satır birleşimi (unionrow) komutu ile birleştirerek tek bir sorgu gibi çalıştırılması sağlanır. Bu durumda oluşacak yeni sorgu

```
'Select { [LabTetikik].[Tüm LabTetikikler].[Testosteron, serbest] } on columns, { [Hasta_Durum].members } on rows From Lab_Islemleri
```

### unionrow

```
Select { [LabTetikik].[Tüm LabTetikikler].[Testosteron, total] } on columns, { [Hasta_Durum].members } on rows From Lab_Islemleri'
```

şeklindedir. Bu sorguya ait sonuç kümesi Şekil 4.39 ile gösterilmiştir. Her iki sorgu sonucu yan yana gelecek şekilde grid bileşenine eklenmiştir.

|         | Testosteron, serbest |
|---------|----------------------|
| Tüm     | 329,00               |
| Ayaktan | 325,00               |
| Yatan   | 4,00                 |

Şekil 4.37 Birinci Sorguya Ait Sonuç Kümesi

|         | Testosteron, total |
|---------|--------------------|
| Tüm     | 340,00             |
| Ayaktan | 331,00             |
| Yatan   | 9,00               |

Şekil 4.38 İkinci Sorguya Ait Sonuç Kümesi

|           | Testosteron, serbest | Testosteron, total |
|-----------|----------------------|--------------------|
| Tüm Hasta | 329,00               | 340,00             |
| Ayaktan   | 325,00               | 331,00             |
| Yatan     | 4,00                 | 9,00               |

Şekil 4.39 Satır Birleşimi Yapılmış Sorguya Ait Sonuç Kümesi

- Satır Birleşimi (UnionRow) ve Sütun Birleşimi (UnionColumn) komutlarının birlikte kullanımı

Hem satır hem de sütun birleşimi komutları bir arada istenilen sayıda kullanılabilir. Dikkat edilmesi gereken husus, öncelikle satır birleşimi ile başlanmasının gerekliliğidir. Yani öncelik, soldan sağa daha sonra yukarıdan

aşağıya şekilde sorgu sonuçları üretilmektedir. Ayrıca, hazırlanan sorguların satır ya da sütun birleşimine göre kesişim içerisinde olduğu diğer sorgularla uygunluk içerisinde olmalıdır. Örneğin, 6 tane ÇBİ cümleciğimiz olsun. Bunlardan ilk 3 tanesi satır birleşimi ile birbirine bağlansın. Geriye kalan 3 tanesi de sütun birleşimi ile bir üst cümleciklere ve satır birleşimi ile birbirlerine bağlansın. Bu örnek için grid bileşeni bünyesinde 6 tane blok bulunacaktır. Bu blokları Şekil 4.40 ile gösterildiği gibi A,B,C,D,E ve F olarak adlandıralım. A,B ve C bloklarını kapsayan ÇBİ cümlecikleri satır birleşimi ile birbirine bağlıdır ve satır başlıkları aynıdır. D,E ve F bloklarını kapsayan ÇBİ cümlecikleri satır birleşimi ile birbirine bağlıdır ve satır başlıkları aynıdır. Aynı zamanda D bloğuna ait ÇBİ cümleciği A bloğuna ait ÇBİ cümleciği ile sütun birleşimi ile birbirine bağlıdır. Benzer şekilde E bloğu B bloğu ile ve F bloğu C bloğu ile sütun birleşimi yapmaktadır. Birleştirilecek olan ÇBİ cümlecikleri öncelikle satır birleşimi ile daha sonra sütun birleşimi ile bağlanmaktadır.

|   |   |   |
|---|---|---|
| A | B | C |
| D | E | F |

Şekil 4.40 Satır ve Sütun Birleşimlerinin Birlikte Kullanımı

Satır ve sütun birleşimini içeren, 'Laboratuar İşlemleri(Lab\_Islemleri)' küpü üzerinde ÇBİ oluşturalım. Birinci sorgu, laboratuar tetkiklerinden 'Serbest Testesteron' işlemini sütun bazında ve hasta durumlarını satır bazında oluştursun. İkinci sorgu, laboratuar tetkiklerinden 'Toplam Testesteron' işlemini sütun bazında ve hasta durumlarını satır bazında oluştursun. Üçüncü sorgu, laboratuar tetkiklerinden 'Serbest Testesteron' işlemini sütun bazında ve 2004 yılına ait tarih bilgilerini satır bazında oluştursun. Dördüncü sorgu, laboratuar tetkiklerinden 'ToplamTestesteron' işlemini sütun bazında ve 2004 yılına ait tarih bilgilerini satır bazında oluştursun. Bu sorguları içeren ÇBİ cümleciği;



'Select { [LabTetikik].[Tüm LabTetikikler].[Testosteron, serbest] } on columns, { [Hasta\_Durum].members } on rows From Lab\_Islemleri

#### **unionrow**

Select { [LabTetikik].[Tüm LabTetikikler].[Testosteron, total] } on columns, { [Hasta\_Durum].members } on rows From Lab\_Islemleri

#### **unioncolumn**

Select { [LabTetikik].[Tüm LabTetikikler].[Testosteron, serbest] } on columns, { [Tarihler].[2004].[Ocak] : [Tarihler].[2004].[Aralık] } on rows From Lab\_Islemleri

#### **unionrow**

Select { [LabTetikik].[Tüm LabTetikikler].[Testosteron, total] } on columns, { [Tarihler].[2004].[Ocak] : [Tarihler].[2004].[Aralık] } on rows From Lab\_Islemleri'

şeklindedir.

İlk iki ÇBİ cümlecği satır birleşimi ile birbirine bağlanmıştır. Üçüncü ÇBİ cümlecği birinci ÇBİ cümlecği ile sütun birleşimi ve dördüncü ÇBİ cümlecği ile satır birleşimi ile bağlanmıştır. Dördüncü ÇBİ cümlecği, ikinci ÇBİ cümlecği ile sütun birleşimi ve üçüncü ÇBİ cümlecği ile satır birleşimi yapmıştır. Yeni oluşturulan ÇBİ sorgu sonucu Şekil 4.41 ile gösterilmiştir.

|   |           | A                    | B                  |  |
|---|-----------|----------------------|--------------------|--|
|   |           | Testosteron, serbest | Testosteron, total |  |
| A | Tüm Hasta | 329,00               | 340,00             |  |
|   | Ayaktan   | 325,00               | 331,00             |  |
|   | Yatan     | 4,00                 | 9,00               |  |
|   | Ocak      | 9,00                 | 17,00              |  |
|   | Şubat     | 19,00                | 9,00               |  |
|   | Mart      | 48,00                | 32,00              |  |
|   | Nisan     | 37,00                | 30,00              |  |
|   | Mayıs     | 22,00                | 25,00              |  |
|   | Haziran   | 38,00                | 24,00              |  |
| C | Temmuz    | 27,00                | 44,00              |  |
|   | Ağustos   | 20,00                | 27,00              |  |
|   | Eylül     | 23,00                | 32,00              |  |
|   | Ekim      | 29,00                | 38,00              |  |
|   | Kasım     | 23,00                | 26,00              |  |
|   | Aralık    | 34,00                | 36,00              |  |
|   |           |                      |                    |  |
|   |           |                      |                    |  |
|   |           | C                    | D                  |  |

Şekil 4.41 4 ÇBİ Cümleciğinden Oluşan Örnek Sorgu Sonucu (A : Birinci Sorgu Sonucu, B : İkinci Sorgu Sonucu, C : Üçüncü Sorgu Sonucu, D : Dördüncü Sorgu Sonucu)

#### 4.4.3.3. Grid Bileşeni Üzerinde Sorgunun Çalıştırılması

Raporlar bir ya da birden çok sorgunun yer aldığı parçalar bütünüdür. Bu bağlamda hazırlanan raporlama aracı çeşitli sorguların üretilmesine imkân tanımalı ve bu sorguların ön izlenmesine izin vermelidir. Kullanıcının sorguda görmek istediği satır, sütun ve hücre değerlerinin adım adım temin edilerek geliştirdiğim grid bileşeni üzerinde izlenmesi mümkün kılınmıştır. Hazırladığım grid arayüzü Microsoft ürün ailesinde yer alan 'Hierarchical Grid' bileşeninin ek özelliklerle genişletilmesiyle oluşturulmuştur. Kullanıcının oluşturduğu ÇBİ cümleciği çalıştırılarak grid üzerinde gösterilir. Bu aşama mevcut bileşen üzerine elde edilen verilerin yazdırılması ile sağlanır. Sıradüzensel bir yapıya sahip sorgularda verilerin en üst seviyedeki veri grubuna göre toparlanması sağlanmıştır. Kullanıcı ihtiyaca göre verinin daha detaylı alt seviyelerine erişebilir. Bu işlem için grid

bileşeninde satır ya da sütün üzerinde grubu ifade eden açıklamanın çift tıklanması gerekmektedir. Bir satır ya da sütün açıklamasının detay bilgisinin var olması ilgili açıklamanın başında yer alan '+' işleci ile kolaylıkla anlaşılabilir. Veri grupları ağaç yapısında olduğu gibi genişletilip daraltılabilmektedir. Genişletilen veri grubuna ait açıklama alanı '-' işlemine sahip olur ve bu aşamada yapılan çift tıklama ile veri grubu daraltılır. Genişletme ve daraltma işlemlerinde her bir aşamada çalıştırılacak ÇBİ cümlecği yeniden üretilir ve bu cümlecğin çalıştırılması ile elde edilen veriler grid bileşeninde uygun hücrelere yerleştirilir. Örneğin;

```
'Select { [Tarihler].[2004].[Ocak] : [Tarihler].[2004].[Aralık] } on columns, { [Ameliyat_Tur].levels(0).members , [Ameliyat_Tur].levels(1).members } on rows From Ameliyat '
```

şeklinde hazırlanmış bir ÇBİ cümlecğinin grid bileşeni üzerindeki görüntüsü Şekil 4.42 ile gösterilmiştir.

|                                      | Mayıs | Haziran | Temmuz | Ağustos | Eylül | Ekim | Kasım | Aralık |
|--------------------------------------|-------|---------|--------|---------|-------|------|-------|--------|
| All Ameliyat_Tur                     |       |         |        |         | 14,00 | 6,00 | 8,00  | 6,00   |
| + 3-ALGOLOJİ-AĞRI TEDAVİSİ           |       |         |        |         |       |      |       |        |
| + ARTRODEZ                           |       |         |        |         |       |      |       |        |
| + ÇOCUK CERRAHİSİ[5]                 |       |         |        |         | 2,00  |      |       |        |
| + DAMAR CERRAHİSİ                    |       |         |        |         |       |      |       |        |
| + GENEL CERRAHİ                      |       |         |        |         | 2,00  |      | 2,00  | 4,00   |
| + GİRİŞİMSEL RADYOLOJİK İŞLEMLER[40] |       |         |        |         |       |      |       |        |
| + GÖĞÜS CERRAHİSİ                    |       |         |        |         |       |      |       |        |
| + GÖZ HASTALIKLARI                   |       |         |        |         |       |      |       |        |
| + İÇ HASTALIKLARI                    |       |         |        |         |       |      |       |        |
| + KADIN HASTALIKLARI VE DOĞUM        |       |         |        |         | 6,00  | 2,00 |       |        |
| + KALP VE DAMAR CERRAHİSİ            |       |         |        |         |       |      | 2,00  |        |

Şekil 4.42 Grid Bileşeni Görünümü

Ameliyat türlerinin yer aldığı satırlar, türlere göre sıradüzensel olarak gruplanmıştır. 'Genel Cerrahi' ana grubu altında yer alan alt grupları incelemek için ilgili satıra gelerek çift tıklamak gerekmektedir. Bu işlem sonrasında ÇBİ cümlecği yeniden hazırlanarak ilgili satırın alt dallarını getirecek şekilde çalıştırılmaktadır ve grid bileşenine yeni satırlar eklenerek detaylandırılma yapılmaktadır. Oluşturulan yeni cümlecik

```
'Select { [Tarihler].[2004].[Ocak] : [Tarihler].[2004].[Aralık] } on columns, { [Ameliyat_Tur].[GENEL CERRAHİ].children} on rows From Ameliyat'
```

şeklindedir.

Bu sorgu kapsamında sütun içerikleri değişken değildir dolayısıyla yukarıda oluşturulmuş yeni ÇBİ ifadesi yeterlidir. Fakat bazı durumlarda her bir hücre için yeni cümlecikler üretmek gerekmektedir. Örneğin;

```
'Select { [Tarihler].levels(0).members , [Tarihler].levels(1).members } on
columns, { [Ameliyat_Tur].levels(0).members , [Ameliyat_Tur].levels(1).members }
on rows from Ameliyat '
```

ÇBİ cümlecği ile hem satır hem de sütun kapsamında sıradüzensel gösterimler yer almaktadır. Bu sorgunun grid bileşenindeki görüntüsü Şekil 4.43 ile gösterilmiştir.

|                            | All Tarihler | + 2004 | + 2005 |
|----------------------------|--------------|--------|--------|
| All Ameliyat_Tur           | 114,00       | 34,00  | 80,00  |
| + 3-ALGOLOJİ-AĞRI TEDAVİSİ |              |        |        |
| + ARTRODEZ                 |              |        |        |
| + ÇOCUK CERRAHİSİ[5]       | 22,00        | 2,00   | 20,00  |
| + DAMAR CERRAHİSİ          |              |        |        |
| + GENEL CERRAHI            | 8,00         | 8,00   |        |
| + GİRİŞİMSEL RADYOLOJİK    |              |        |        |
| + GÖĞÜS CERRAHİSİ          |              |        |        |
| + GÖZ HASTALIKLARI         |              |        |        |
| + İÇ HASTALIKLARI          |              |        |        |
| + KADIN HASTALIKLARI VE    | 41,00        | 8,00   | 33,00  |
| + KALP VE DAMAR            | 2,00         | 2,00   |        |
| KAN VE LENF SİSTEMLERİ     |              |        |        |
| + KARDİYOLOJİ              |              |        |        |

Şekil 4.43 Grid Bileşeni Görünümü

Öncelikle tarih sütununda yer alan “2005” veri grubunu açalım.

```
'Select { [Tarihler].[2005].children} on columns, { [Ameliyat_Tur].levels(0).members
, [Ameliyat_Tur].levels(1).members } on rows From Ameliyat' şeklinde yeni bir
ÇBİ üretilecektir.
```

Şimdi de ameliyat türleri arasından 'KADIN HASTALIKLARI' veri grubunu açalım. Bu işlem sonrasında her bir sütun kontrol edilerek çeşitli sayıda yeni ÇBİ cümlecği üretilecektir.

```
'Select { [Tarihler].levels(0).members , [Tarihler].levels(1).members } on
columns, { [Ameliyat_Tur].[KADIN HASTALIKLARI VE DOĞUM].children} on
```

rows From Ameliyat' cümlecği ile ilgili ameliyat türünün alt grupları grid bileşenine eklenir. Fakat bu cümlecik ilgili sorgu için yeterli değildir. Çünkü grid bileşeni üzerinde '2005' veri grubu genişletilerek aylara bölünmüştür. Bu durumda '2005-Ocak' açıklaması için 'Select { [Tarihler].[2005].[Ocak] } on columns, { [Ameliyat\_Tur].[KADIN HASTALIKLARI VE DOĞUM].children} on rows From Ameliyat' şeklinde , '2005-Şubat' açıklaması için 'Select { [Tarihler].[2005].[Şubat] } on columns, { [Ameliyat\_Tur].[KADIN HASTALIKLARI VE DOĞUM].children } on rows From Ameliyat' şeklinde yeni ÇBİ cümlecikleri üretilir. Bu işlem her bir sütün değeri kontrol edilerek sırayla yapılır. Yeni oluşan sorguya ait veriler Şekil 4.44 ile gösterilmiştir.

|                               |                           | + 2004 |       | - 2005 |       |      |       |    |
|-------------------------------|---------------------------|--------|-------|--------|-------|------|-------|----|
|                               |                           | 2004   | 2005  | Ocak   | Şubat | Mart | Nisan | Mi |
| + DAMAR CERRAHISI             | DAMAR CERRAHISI           |        |       |        |       |      |       |    |
| + GENEL CERRAHI               | GENEL CERRAHI             | 8,00   |       |        |       |      |       |    |
| + GİRİŞİMSSEL RADYOLOJİK      | GİRİŞİMSSEL RADYOLOJİK    |        |       |        |       |      |       |    |
| + GÖĞÜS CERRAHISI             | GÖĞÜS CERRAHISI           |        |       |        |       |      |       |    |
| + GÖZ HASTALIKLARI            | GÖZ HASTALIKLARI          |        |       |        |       |      |       |    |
| + İÇ HASTALIKLARI             | İÇ HASTALIKLARI           |        |       |        |       |      |       |    |
| - KADIN HASTALIKLARI VE DOĞUM | KADIN HASTALIKLARI VE     | 8,00   | 33,00 | 3,00   |       |      |       |    |
|                               | + ADNEKSLER               |        |       |        |       |      |       |    |
|                               | + ENDOSKOPIK İŞLEMLER(20) |        |       |        |       |      |       |    |
|                               | + FİSTÜL OPERASYONLARI    |        |       |        |       |      |       |    |
|                               | + GEBELİK VE DOĞUM        | 8,00   | 33,00 | 3,00   |       |      |       |    |
|                               | + İNİTİRFERTİLİZASYON     |        |       |        |       |      |       |    |

Şekil 4.44 Grid Bileşeni Görünümü

Daha özet bilgiye erişilmek istendiğinde de "-" işleci ile görüntülenen ve detaylanmış başlıkların çift tıklanarak kapatılması sağlanır. Örneğin "-2005" açıklaması üzerinde çift tıklama yapıldığında 2005 yılına ait detaylı bilgiler kaldırılarak bu yıla ait özet bilgi grid üzerinde görüntülenecektir.

|                               |                           | All Tarihler | + 2004 | + 2005 |
|-------------------------------|---------------------------|--------------|--------|--------|
| All Ameliyat_Tur              |                           | 114,00       | 34,00  | 80,00  |
| + 3-ALGOLOJİ-AĞRI TEDAVİSİ    | 3 ALGOLOJİ AĞRI TEDAVİSİ  |              |        |        |
| + ARTRODEZ                    | ARTRODEZ                  |              |        |        |
| + ÇOCUK CERRAHİSİ[5]          | ÇOCUK CERRAHİSİ[5]        | 22,00        | 2,00   | 20,00  |
| + DAMAR CERRAHİSİ             | DAMAR CERRAHİSİ           |              |        |        |
| + GENEL CERRAHI               | GENEL CERRAHI             | 8,00         | 8,00   |        |
| + GİRİŞİMSEL RADYOLOJİK       | GİRİŞİMSEL RADYOLOJİK     |              |        |        |
| + GÖĞÜS CERRAHİSİ             | GÖĞÜS CERRAHİSİ           |              |        |        |
| + GÖZ HASTALIKLARI            | GÖZ HASTALIKLARI          |              |        |        |
| + İÇ HASTALIKLARI             | İÇ HASTALIKLARI           |              |        |        |
| - KADIN HASTALIKLARI VE DOĞUM | KADIN HASTALIKLARI VE     | 41,00        | 8,00   | 33,00  |
|                               | + ADNEKSLER               |              |        |        |
|                               | + ENDOSKOPİK İŞLEMLER[20] |              |        |        |
|                               | + FİSTÜL OPERASYONLARI    |              |        |        |
|                               | + GEBELİK VE DOĞUM        | 41,00        | 8,00   | 33,00  |

Şekil 4.45 Grid Bileşeni Görünümü

Grid bileşenine eklenmiş bu yeni özellik sayesinde ÇİAİ uygulamalarında yer alması gereken 'roll-up, boyutlar üzerinde gruplama seviyesinin artırılması' ve 'drill-down, daha yüksek ayrıntı düzeyine inilmesi' işlemlerinin gerçekleştirilmesi sağlanmıştır.

#### 4.4.3.4. Sorguların Grafikselleşmeleri

ÇBİ üretilerek grid bileşeni üzerinde sorgu sonuçlarının izlenebildiği sorgular aynı zamanda grafikselleştirme ile de ifade edilebilmektedir. Hazırladığım yazılım kapsamında çubuk (bar), çizgi (line) ve dairesel (pie) çizgilerle aynı ÇBİ ifadesi ile elde edilen sonuç kümesi 3 farklı şekilde grafik ortamında gösterilebilmektedir.

Örnek olarak 'Dogum' küpü üzerinde, 2004 yılına ait Ocak ve Haziran ayları arasında gerçekleşen doğum sayılarının doğum türlerine göre listelenmesini sağlayan bir sorgu hazırlayalım. Sorguya ait ÇBİ cümlecisi aşağıda verilmiştir;

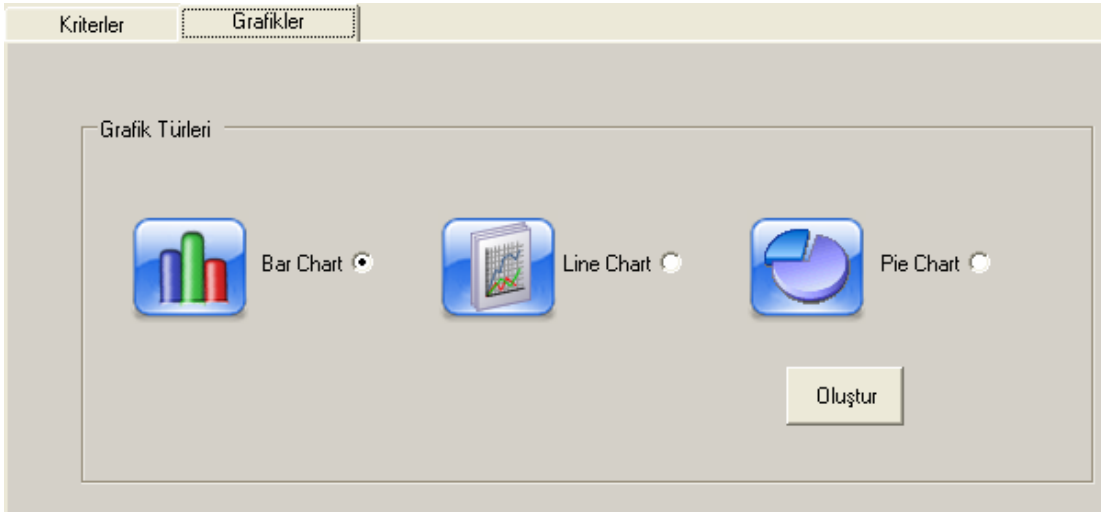
```
'Select { [DogumTuru].members } on columns, { [Tarihler].[2004].[Ocak] : [Tarihler].[2004].[Haziran] } on rows From Dogum'
```

Sorgu sonucunun grid bileşeni ile gösterimi Şekil 4.46 ile gösterilmiştir.

|         | Tüm Doğum Türleri | Normal | Sezeryan |
|---------|-------------------|--------|----------|
| Ocak    | 136,00            | 34,00  | 102,00   |
| Şubat   | 112,00            | 42,00  | 70,00    |
| Mart    | 142,00            | 50,00  | 92,00    |
| Nisan   | 150,00            | 51,00  | 99,00    |
| Mayıs   | 106,00            | 24,00  | 82,00    |
| Haziran | 118,00            | 13,00  | 105,00   |

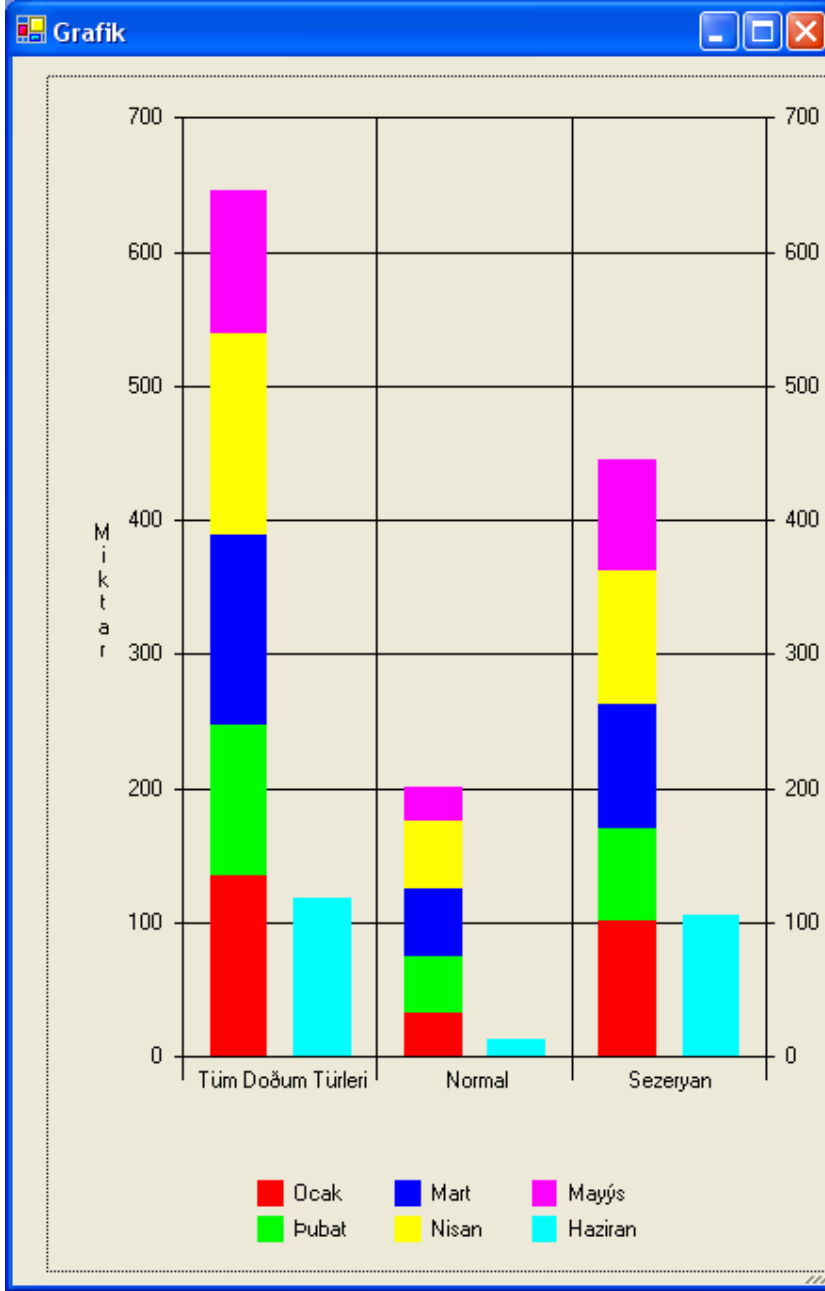
Şekil 4.46 Doğum küpü üzerinde hazırlanmış ÇBİ sorgu sonucu

'Grafikler' sayfasının aktif hale getirilerek çubuk (bar), çizgi (line) ve dairesel (pie) çizge türlerinden birisi seçilerek, grafiğin oluşturulması sağlanır.



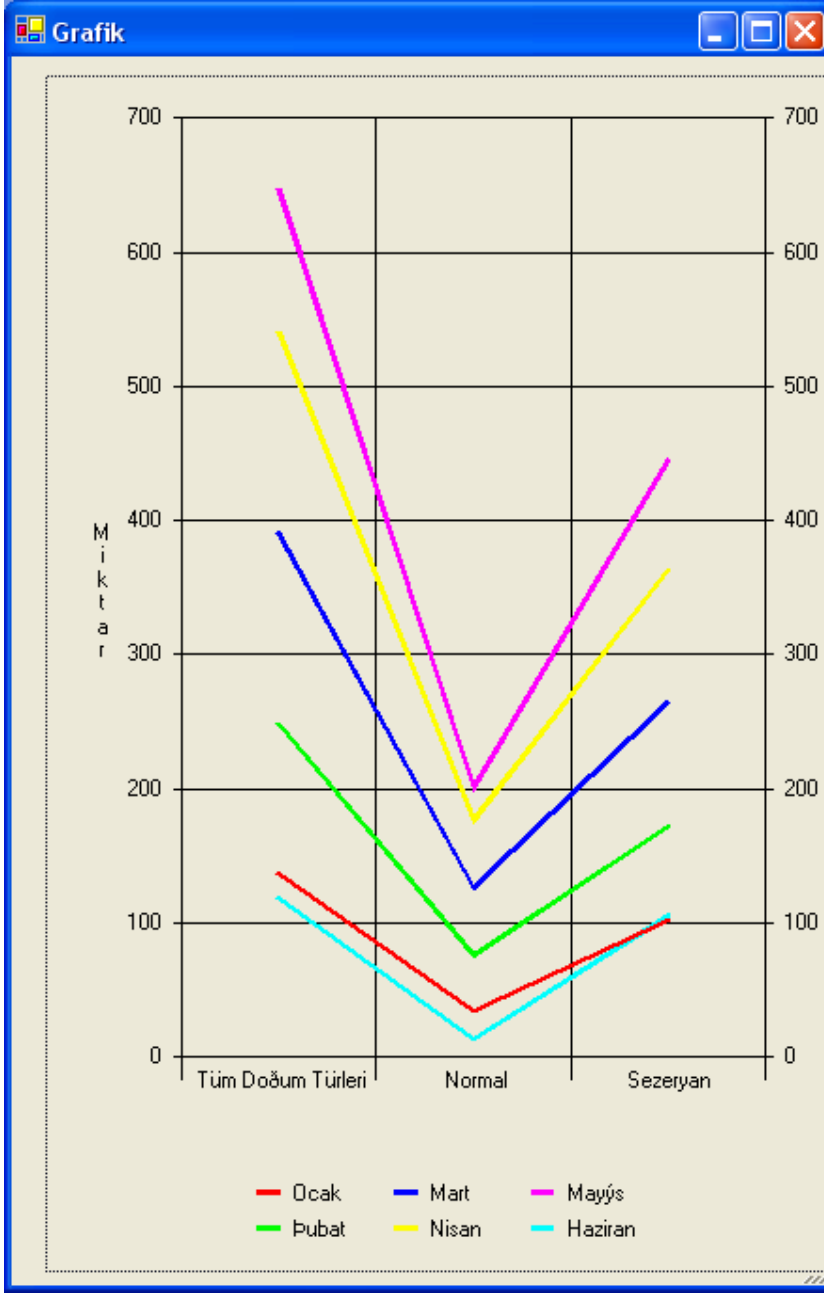
Şekil 4.47 'Grafikler' Sayfa Görünümü

Örnek olarak verdiğimiz sorgunun çubuk, çizgi ve dairesel çizge türlerine göre oluşturulan sonuçları Şekil 4.48, Şekil 4.49 ve Şekil 4.50 ile gösterilmiştir.

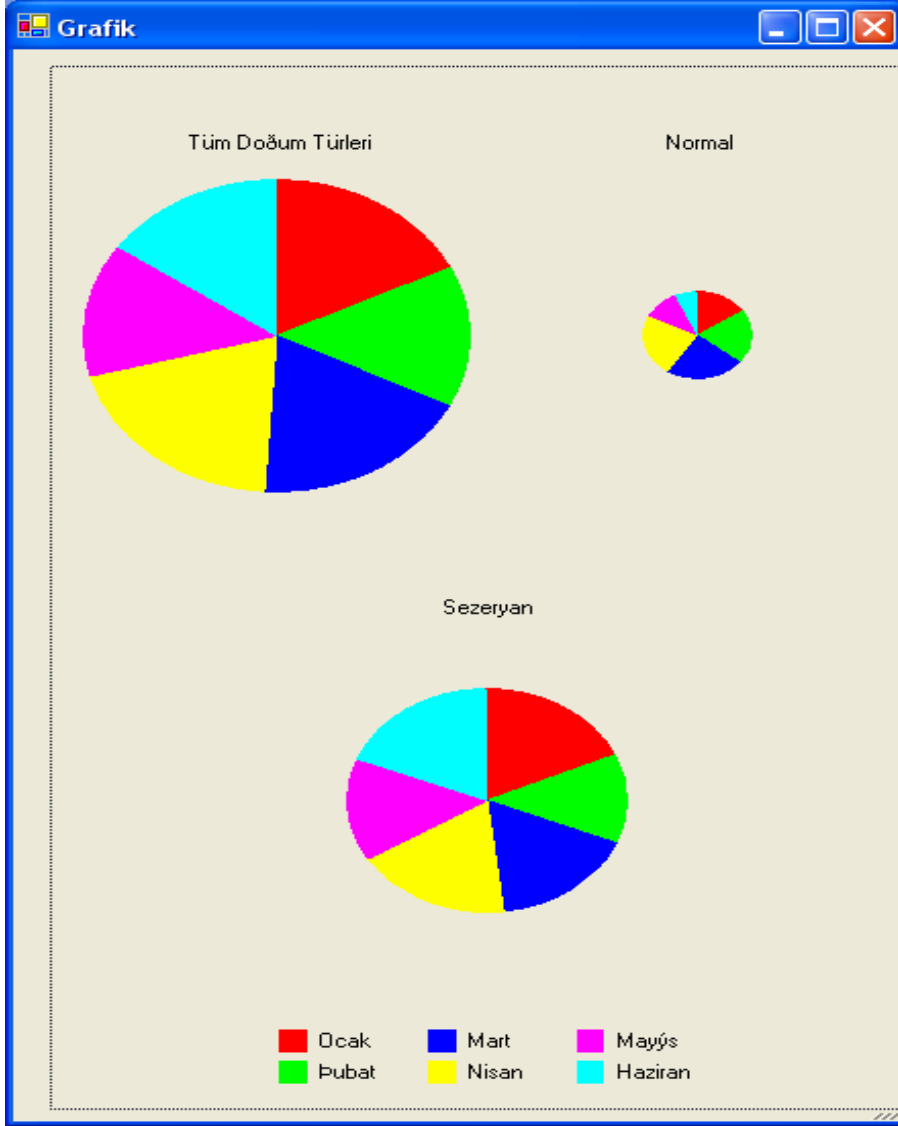


Şekil 4.48 'Doğum' küpü üzerinde hazırlanmış sorgunun çubuk çizge görünümü





Şekil 4.49 'Dogum' küpü üzerinde hazırlanmış sorgunun çizgi çizge görünümü

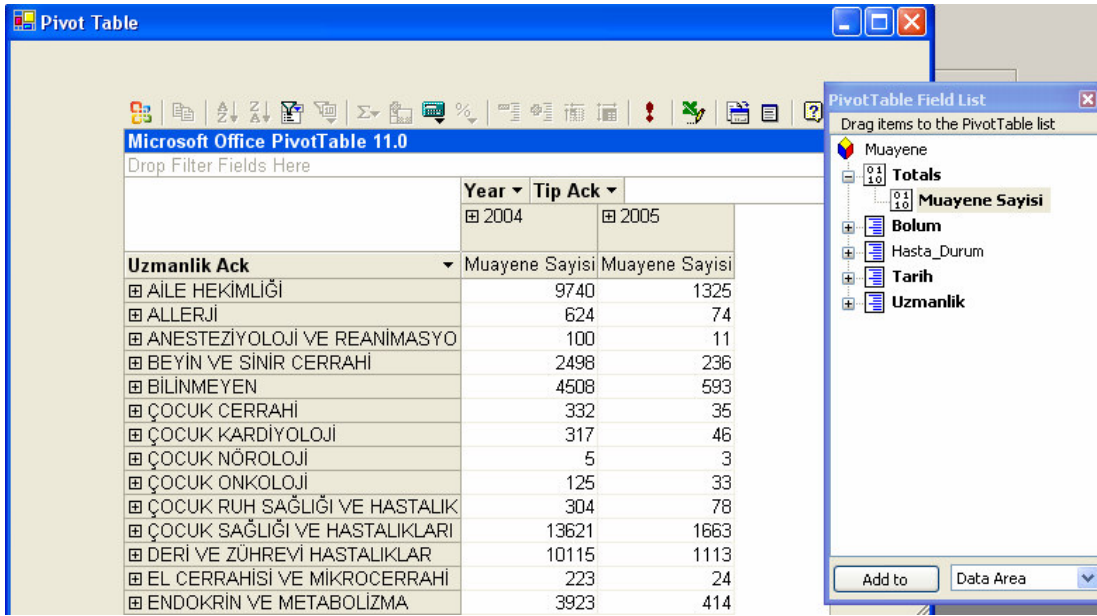


Şekil 4.50 'Dogum' küpü üzerinde hazırlanmış sorgunun dairesel çizge görünümü

#### 4.4.4. 'Pivot Table' Servisi

'Pivot Table' servisi Microsoft ürün ailesinde yer alan ve kullanıcılar için tasarlanmış bir işlem servisidir. Hazırladığım yazılım bu servisi de bünyesinde barındırmaktadır. Dolayısıyla, sürükle bırak yönteminde kolay sorgulama yapılması için bir alternatif sağlanmıştır. 'Pivot Table' servisine ait kullanıcı arayüzü Şekil 4.51 ile gösterilmiştir. Küp isimlerinin yer aldığı listeden bir küp seçtikten sonra 'Diğer İşlemler' menüsünden 'Pivot Table Servisi' menü başlığı

seçilerek çalıştırılır. Seçilmiş olan küpe ait boyut ve ölçüm değerleri sürükleyip bırak yöntemiyle PivotTable arayüzüne bırakılır.



The screenshot shows the Microsoft Office PivotTable 11.0 interface. The main window displays a PivotTable with the following data:

| Uzmanlık Ack                  | Year  | Tip Ack |
|-------------------------------|-------|---------|
|                               | 2004  | 2005    |
| AİLE HEKİMLİĞİ                | 9740  | 1325    |
| ALLERJİ                       | 624   | 74      |
| ANESTEZİYOLOJİ VE REANİMASYON | 100   | 11      |
| BEYİN VE SINIR CERRAHİ        | 2498  | 236     |
| BİLİNMEYEN                    | 4508  | 593     |
| ÇOCUK CERRAHİ                 | 332   | 35      |
| ÇOCUK KARDİYOLOJİ             | 317   | 46      |
| ÇOCUK NÖROLOJİ                | 5     | 3       |
| ÇOCUK ONKOLOJİ                | 125   | 33      |
| ÇOCUK RUH SAĞLIĞI VE HASTALIK | 304   | 78      |
| ÇOCUK SAĞLIĞI VE HASTALIKLARI | 13621 | 1663    |
| DERİ VE ZÜHREVİ HASTALIKLAR   | 10115 | 1113    |
| EL CERRAHİSİ VE MİKROCERRAHİ  | 223   | 24      |
| ENDOKRİN VE METABOLİZMA       | 3923  | 414     |

The PivotTable Field List task pane on the right shows the following fields:

- Muayene
- Totals
- Muayene Sayisi
- Bolum
- Hasta\_Durum
- Tarih
- Uzmanlik

Şekil 4.51 Pivot Table Servis Arayüzü

#### 4.4.5. Küp Editörü

Küp editörü Microsoft Analysis Server ürününde yer alan ve oluşturulmuş küplerin içeriğinin incelenmesini sağlayan bir arayüzdür. Hazırladığım uygulama yazılımı bu arayüzü de kullanıcılara sunmaktadır. Bu sayede, küp editörünü kullanmaya alışkın olan kullanıcılar için hızlı sorgulama imkânı oluşmaktadır. Küp isimlerinin yer aldığı listeden bir küp seçildikten sonra 'Diğer İşlemler' menüsünden 'Küp Editörü' menü başlığı seçilerek çalıştırılır. Seçilmiş olan küpe ait boyutlar sürükleyip bırak yöntemiyle küp editörü arayüzüne bırakılır. Küp editörüne ait kullanıcı arayüzü Şekil 4.52 ile gösterilmiştir.

Küp Editörü

Hasta\_Durum: All Hasta\_Durum

Measures: Muayene Sayisi

Bolum: All Bolum

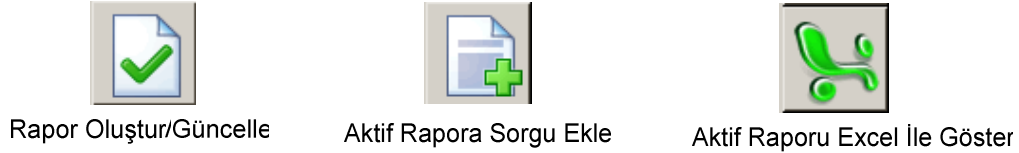
|                         | + Year     |           |
|-------------------------|------------|-----------|
|                         | + 2004     | + 2005    |
| + Uzmanlik Ack          |            |           |
| All Uzmanlik            | 125.089,00 | 14.318,00 |
| + ADLİ TIP              |            |           |
| + AİLE HEKİMLİĞİ        | 9.740,00   | 1.325,00  |
| + ALLERJİ               | 624,00     | 74,00     |
| + ANESTEZİYOLOJİ VE RE  | 100,00     | 11,00     |
| + BEYİN VE SİNİR CERRAI | 2.498,00   | 236,00    |
| + BİLİNMEYEN            | 4.508,00   | 593,00    |
| + BİYOKİMYA VE KLİNİK B |            |           |
| + ÇOCUK CERRAHI         | 332,00     | 35,00     |
| + ÇOCUK KARDİYOLOJİ     | 317,00     | 46,00     |
| + ÇOCUK NÖROLOJİ        | 5,00       | 3,00      |
| + ÇOCUK ONKOLOJİ        | 125,00     | 33,00     |
| + ÇOCUK RUH SAĞLIĞI VE  | 304,00     | 78,00     |
| + ÇOCUK SAĞLIĞI VE HAST | 13.621,00  | 1.663,00  |
| + DERİ VE ZÜHREVİ HAST  | 10.115,00  | 1.113,00  |
| + DIŞ HASTALIKLARI VE T |            |           |

Şekil 4.52 Muayene Küp İçeriğini Gösteren Küp Editörü Arayüzü

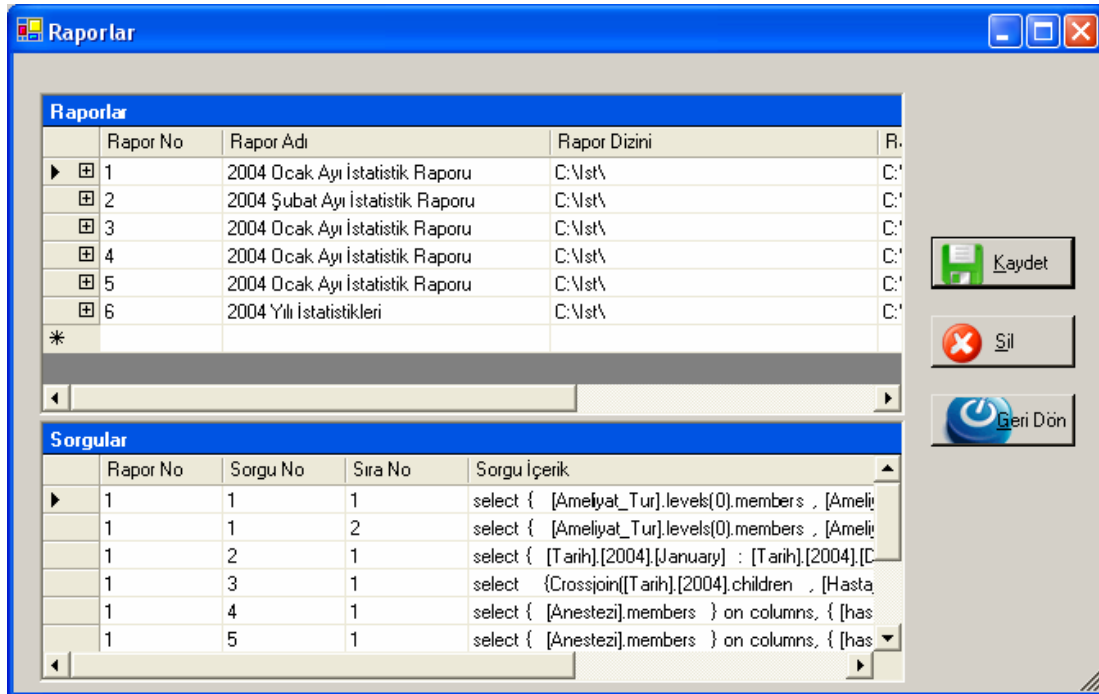
#### 4.4.6. Sorgulardan Rapor Tanımlama

Bir ya da birden çok ÇBİ cümlecığının çalıştırılması ile oluşan raporlar tanımlanabilmektedir. Öncelikle, istenilen raporun hangi ÇBİ 'leri kapsayacağı tanımlanmalıdır. Bu nedenle, ilişkiel veri tabanı üzerinde 'Raporlar ve Sorgular' olmak üzere iki tablo bulunmaktadır. 'Raporlar' tablosu, 'Rapor Numarası, Rapor Adı, Raporun Oluşturulacağı Dizin' bilgilerini içermektedir. 'Sorgular' tablosu, 'Rapor Numarası, Sorgu Numarası, Sıra Numarası, Sorguya Ait ÇBİ İçeriği' bilgilerini içermektedir. 'Rapor Numarası', 'Raporlar' tablosunda tanımlı olan rapora ait yabancı anahtar alanıdır. 'Sorgu Numarası', ilgili raporun içerisinde bulunan sorguların numarasıdır. 'Sıra Numarası', belirtilen raporun belirtilen sorgu numarasındaki sıra bilgisidir. Raporlar, öncelikle sorgu numarasına göre daha sonra sıra numaralarına göre sorguları kapsamaktadır. Aynı sorgu numarasına ait sorgular, rapor ortamında yan yana üretilecek ve aynı satırları paylaşan sorgulardır. Örneğin sorgu numarası 1, sıra numarası 1 ve 2 olmak üzere iki adet sorguyu ele alalım. Öncelikle sıra numarası 1 olan sorgu çalıştırılarak sonuç

kümesi bir Excel sayfası üzerinde işlenecektir. Daha sonra sıra numarası 2 olan sorgu çalıştırılacak ve ilk sorgunun bittiği sütun değerlerinden başlamak üzere ve her iki sorgu yan yana görüntülenecek şekilde çıktı üretecektir. Raporlar, hazırladığım yazılım kapsamında oluşturulmakta ve grid bileşeninde gösterilen her bir sorgu cümlecği aktif durumda bulunan rapora yeni bir sorgu olarak eklenebilmektedir. Bu işlem için, 'Rapor Oluştur/Güncelle' ve 'Aktif Rapora Sorgu Ekle' düğmeleri kullanılmaktadır. Rapor ve sorgulara ait veriler Şekil 4.54 ile gösterilen izleme ekranı ile değiştirilebilmektedir.



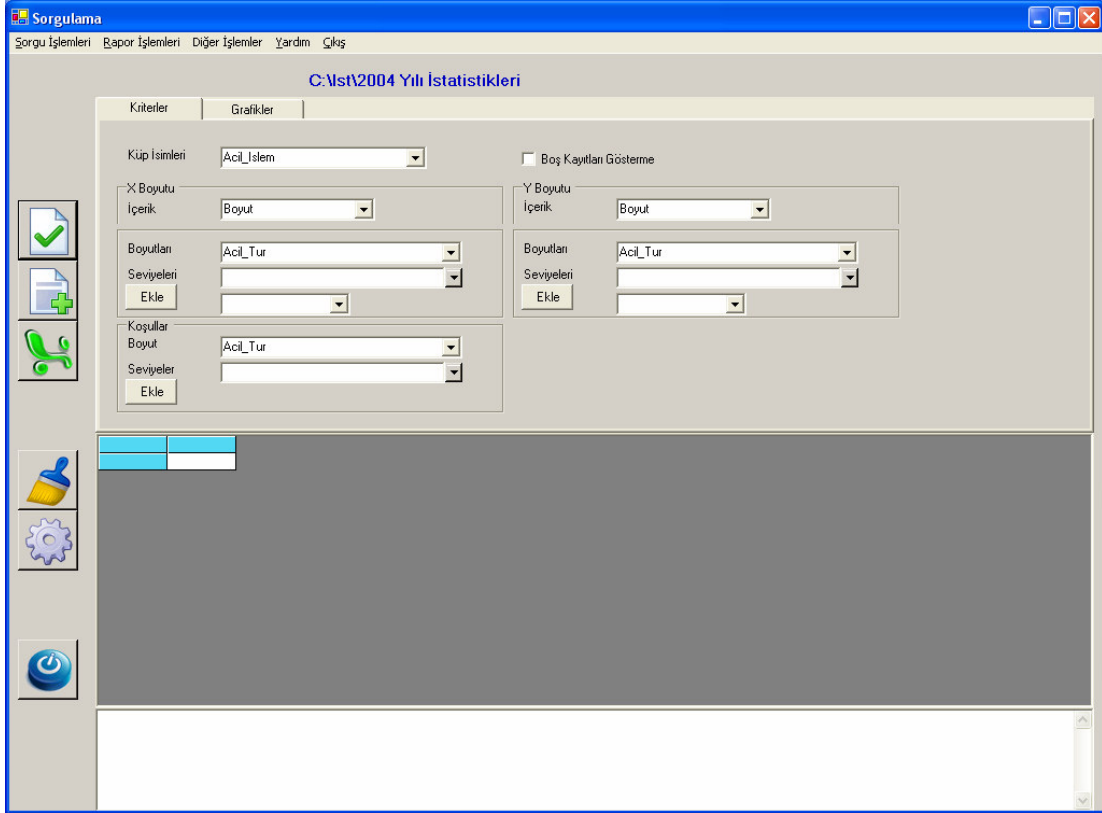
Şekil 4.53 Raporlama İçin Kullanılan Düğmeler



Şekil 4.54 Rapor ve Sorgu İzleme Formu

#### 4.4.7. Rapor Sonucunun Microsoft Excel Ürünü İle Gösterilmesi

'Rapor Oluştur/Güncelle' düğmesine basılarak sistemde tanımlı olan raporların ve raporları oluşturan sorguların incelendiği izleme formuna erişilmektedir. Bu form üzerinde yeni bir rapor oluşturulabilir ya da mevcut bir rapor seçilerek ana ekrana geri dönüş sağlanır. Bu işlem ile seçilmiş olan raporun aktif hale gelmesi sağlanır. Aktif duruma getirilmiş olan rapora 'Aktif Raporu Sorgu Ekle' düğmesi kullanılarak grid bileşeni üzerinde görüntülenen ÇBİ cümlecikleri yeni sorgu olarak eklenebilir. Hangi raporun aktif olduğu ana ekranın üzerinde yer alan rapor ismi ile takip edilebilir. Örneğin; '2004 Yılı İstatistikleri' adlı rapor aktif ise Şekil 4.55 görüldüğü gibi bir bilgilendirme ekran üzerinde yer almaktadır.



Şekil 4.55 " 2004 Yılı İstatistikleri" Adlı Raporun Aktif Rapor Olarak Seçilmesi Durumu

'Aktif Raporu Excel İle Göster' düğmesinin seçilmesi ile veri tabanı üzerinde bulunan 'Sorgular' tablosunda yer alan aktif rapora ait sorgular YSD ile alınarak raporu oluşturan sorgu cümlecikleri elde edilmektedir. Sorgu numaralarına göre

sırayla her bir sorgunun çalıştırılarak Microsoft Excel ürünü üzerinde uygun hücrelere sorgu sonuçlarının yazılması sağlanmaktadır. Sorgu ve sıra numaraları sırasına göre sorgular çalıştırılmaktadır. Grid bileşeninde yer alan satır ve sütun başlıkları Excel sayfası üzerinde mavi, koyu ve yatık olarak yazdırılmaktadır. Her bir satır ve sütunun kesişim noktası ilgili sorgunun sayılabilir ölçüm değerlerini veren içerikle doldurulmaktadır. Yeni bir sorguya geçiş esnasında düz bir çizgi ile ayırım belirtilmektedir. Yeni sıra numarasına sahip bir kayda gelindiğinde, bir önceki sorgunun bittiği sütundan itibaren yeni sorgu Excel sayfasına işlenmektedir. Sorgu numarası değiştiğinde bir önceki sorgunun bittiği satırdan itibaren yeni sorgu sonuçları Excel sayfasına işlenmektedir. Örnek bir rapor içeriği Şekil 4.56 ile gösterilmiştir.

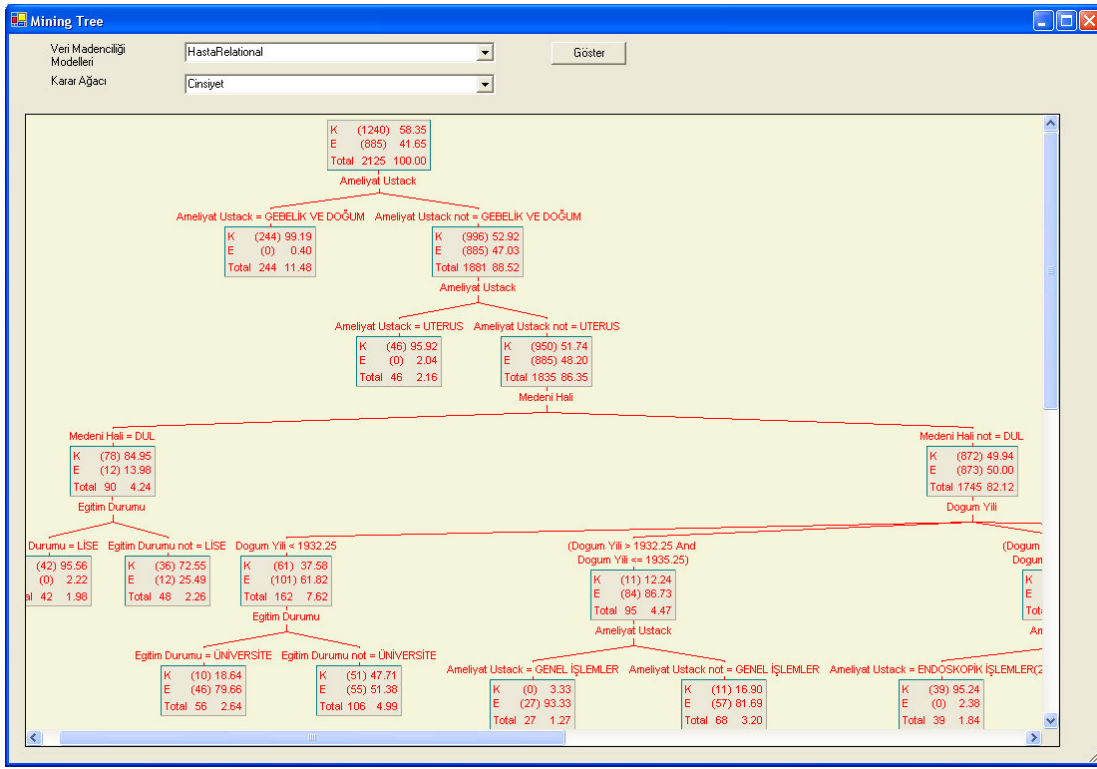
|     | A                                   | B        | C        | D         | E        | F        | G        | H        | I        | J        | K        | L        | M         | N | O |
|-----|-------------------------------------|----------|----------|-----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|---|---|
| 88  | <b>İŞ VE MESLEK HASTALIKLARI</b>    |          |          |           |          |          |          |          |          |          |          |          |           |   |   |
| 89  | <b>EL CERRAHİSİ VE MIKROCERRAHİ</b> |          |          |           |          |          |          |          |          |          |          |          |           |   |   |
| 90  |                                     |          |          |           |          |          |          |          |          |          |          |          |           |   |   |
| 91  |                                     | Ocak     | Şubat    | Mart      | Nisan    | Mayıs    | Haziran  | Temmuz   | Ağustos  | Eylül    | Ekim     | Kasım    | Aralık    |   |   |
| 92  |                                     | 209,00   | 209,00   | 209,00    | 209,00   | 209,00   | 209,00   | 209,00   | 209,00   | 209,00   | 209,00   | 209,00   | 209,00    |   |   |
| 93  |                                     | 8,00     | 8,00     | 8,00      | 8,00     | 8,00     | 8,00     | 8,00     | 8,00     | 8,00     | 8,00     | 8,00     | 8,00      |   |   |
| 94  |                                     | 123,00   | 123,00   | 123,00    | 123,00   | 123,00   | 123,00   | 123,00   | 123,00   | 123,00   | 123,00   | 123,00   | 123,00    |   |   |
| 95  |                                     | 78,00    | 78,00    | 78,00     | 78,00    | 78,00    | 78,00    | 78,00    | 78,00    | 78,00    | 78,00    | 78,00    | 78,00     |   |   |
| 96  |                                     |          |          |           |          |          |          |          |          |          |          |          |           |   |   |
| 97  |                                     | Ocak     | Şubat    | Mart      | Nisan    | Mayıs    | Haziran  | Temmuz   | Ağustos  | Eylül    | Ekim     | Kasım    | Aralık    |   |   |
| 98  |                                     | 5.164,00 | 3.446,00 | 4.248,00  | 4.272,00 | 4.436,00 | 4.262,00 |          |          |          |          |          |           |   |   |
| 99  |                                     |          |          |           |          |          |          |          |          |          |          |          |           |   |   |
| 100 |                                     | Ocak     | Şubat    | Mart      | Nisan    | Mayıs    | Haziran  | Temmuz   | Ağustos  | Eylül    | Ekim     | Kasım    | Aralık    |   |   |
| 101 |                                     | 166,58   | 118,83   | 137,03    | 142,40   | 143,10   | 142,07   |          |          |          |          |          |           |   |   |
| 102 |                                     |          |          |           |          |          |          |          |          |          |          |          |           |   |   |
| 103 |                                     | Ocak     | Şubat    | Mart      | Nisan    | Mayıs    | Haziran  | Temmuz   | Ağustos  | Eylül    | Ekim     | Kasım    | Aralık    |   |   |
| 104 |                                     | 79,70    | 56,86    | 65,57     | 68,13    | 68,47    | 67,97    |          |          |          |          |          |           |   |   |
| 105 |                                     |          |          |           |          |          |          |          |          |          |          |          |           |   |   |
| 106 |                                     | Ocak     | Şubat    | Mart      | Nisan    | Mayıs    | Haziran  | Temmuz   | Ağustos  | Eylül    | Ekim     | Kasım    | Aralık    |   |   |
| 107 |                                     | 3,94     | 3,74     | 3,30      | 3,19     | 3,62     | 3,73     |          |          |          |          |          |           |   |   |
| 108 |                                     |          |          |           |          |          |          |          |          |          |          |          |           |   |   |
| 109 |                                     | Ocak     | Şubat    | Mart      | Nisan    | Mayıs    | Haziran  | Temmuz   | Ağustos  | Eylül    | Ekim     | Kasım    | Aralık    |   |   |
| 110 |                                     | 136,00   | 112,00   | 142,00    | 150,00   | 106,00   | 118,00   | 127,00   | 114,00   | 115,00   | 88,00    | 125,00   | 126,00    |   |   |
| 111 |                                     | 34,00    | 42,00    | 50,00     | 51,00    | 24,00    | 13,00    | 17,00    | 24,00    | 18,00    | 19,00    | 22,00    | 23,00     |   |   |
| 112 |                                     | 102,00   | 70,00    | 92,00     | 99,00    | 82,00    | 105,00   | 110,00   | 90,00    | 97,00    | 69,00    | 103,00   | 103,00    |   |   |
| 113 |                                     |          |          |           |          |          |          |          |          |          |          |          |           |   |   |
| 114 |                                     | Ocak     | Şubat    | Mart      | Nisan    | Mayıs    | Haziran  | Temmuz   | Ağustos  | Eylül    | Ekim     | Kasım    | Aralık    |   |   |
| 115 |                                     | 8.661,00 | 7.312,00 | 10.063,00 | 9.867,00 | 9.534,00 | 9.175,00 | 9.483,00 | 8.989,00 | 9.862,00 | 9.202,00 | 9.501,00 | 11.743,00 |   |   |
| 116 |                                     |          |          |           |          |          |          |          |          |          |          |          |           |   |   |
| 117 |                                     | 689,00   | 653,00   | 840,00    | 810,00   | 808,00   | 757,00   | 920,00   | 671,00   | 876,00   | 841,00   | 781,00   | 1.061,00  |   |   |
| 118 |                                     | 36,00    | 31,00    | 37,00     | 76,00    | 77,00    | 40,00    | 57,00    | 36,00    | 63,00    | 33,00    | 56,00    | 65,00     |   |   |
| 119 |                                     | 9,00     | 3,00     | 6,00      | 4,00     | 4,00     | 1,00     | 1,00     | 1,00     | 1,00     | 2,00     | 8,00     |           |   |   |
| 120 |                                     | 170,00   | 148,00   | 196,00    | 220,00   | 199,00   | 184,00   | 195,00   | 166,00   | 194,00   | 207,00   | 179,00   | 220,00    |   |   |
| 121 |                                     | 174,00   | 171,00   | 313,00    | 341,00   | 405,00   | 335,00   | 363,00   | 304,00   | 333,00   | 381,00   | 425,00   | 549,00    |   |   |
| 122 |                                     |          |          |           |          |          |          |          |          |          |          |          |           |   |   |
| 123 |                                     | 17,00    | 10,00    | 28,00     | 24,00    | 11,00    | 33,00    | 21,00    | 27,00    | 28,00    | 19,00    | 25,00    | 24,00     |   |   |
| 124 |                                     | 22,00    | 16,00    | 36,00     | 22,00    | 21,00    | 22,00    | 21,00    | 6,00     | 29,00    | 16,00    | 12,00    | 22,00     |   |   |
| 125 |                                     |          |          |           |          |          |          |          |          |          |          |          |           |   |   |

Şekil 4.56 Excel üzerinde hazırlanmış rapor çıktısı

#### 4.4.8. Karar Ağacı Madencilik Modelinin Çizimsel Gösterimi

Karar ağacı algoritmasını uygulayarak geliştirilen veri madenciliği modellerinin, hazırlanan uygulamayı yazılımı ile karar destek amaçlı kullanılması sağlanmıştır.

'Diğer İşlemler' menüsünden 'Veri Madenciliği' menü başlığı seçilerek yeni bir arayüze geçiş sağlanır. Bu arayüz üzerinde, bağlantı cümlecisi ile erişilen sistemde oluşturulmuş veri madenciliği modellerinin ve her model için tanımlı karar ağaçlarının listelenmesi sağlanmıştır. 'Veri Madenciliği Modelleri' başlıklı liste üzerinden istenilen madencilik modelinin seçilmesi ile 'Karar Ağacı' başlıklı listenin seçilen madencilik modelinde yer alan karar ağaçlarını ifade etmesi sağlanır. Şekil 4.57 üzerinde veri ambarında bulunan ilişkisel tablolar kullanılarak hazırlanmış veri madenciliği modelinin 'Cinsiyet' karar ağacını ifade eden şekilsel gösterimi yer almaktadır.



Şekil 4.57 'Cinsiyet' Karar Ağacının Görünümü

#### 4.4.9. Madencilik Modelinin Sanal Küplerle Sorgulanması

ÇİAİ küplerinin kaynak olarak kullanıldığı veri madenciliği modellerinde elde edilen sonuç kümesi sanal boyutlar ve sanal küpler üzerinde saklanabilmektedir. Elde edilen sanal küpler, ÇBİ cümlecikleri ile sorgulanarak grid bileşeni üzerinde ya da tanımlanacak bir rapor kapsamında Excel sayfaları ile incelenebilmektedir.



Örneğin, hasta bilgileri küpü üzerinden üretilmiş olan HastaBlgOLAP isimli veri madenciliği modeline ait sonuç kümesi HastaBlgVC isimli sanal küp ile saklanmıştır. Bu küp üzerinden yapılacak sorgulamalarla, madencilik kapsamında elde edilen sonuçların izlenmesi sağlanabilmektedir. Örneğin;

```
'Select { [Cinsiyet].members , [Cinsiyet].[Cinsiyet%E] , [Cinsiyet].[Cinsiyet%K] }
on          columns,          {          [HastaBlgDim].levels(0).members,
[HastaBlgDim].levels(1).members } on rows From HastaBlgVC'
```

şeklinde hazırlanmış bir ÇBİ cümlecği ile veri madenciliği modeli uygulanarak üretilmiş 'HastaBlgDim' adlı sanal boyut satırlarda ve 'Cinsiyet' boyutuna ait üyeler ve hesaplanmış öge olarak tanımlanmış 'Cinsiyet%K ve Cinsiyet%E' bilgi alanları sütunlarda olacak şekilde sorgu sonucu üretilecektir. Elde edilen çıktı görüntüsü Şekil 4.58 ile gösterilmiştir. Grid bileşeni üzerinde, fare ile çift tıklayarak sanal boyut üzerinde bulunan madencilik sonuçları detaylı olarak incelenebilmektedir.

|   | All Cinsiyet | E   | K     | Cinsiyet%E | Cinsiyet%K |
|---|--------------|-----|-------|------------|------------|
| All   | 2.125        | 885 | 1.240 | 42         | 58         |
| Hasta_Numara.Hasta No. Ameliyat Ustack = GEBELİK VE       | 244          |     | 244   |            | 100        |
| + Hasta_Numara.Hasta No. Ameliyat Ustack not = GEBELİK VE | 1.881        | 885 | 996   | 47         | 53         |

Şekil 4.58 Sanal küp üzerinde çalıştırılan ÇBİ sonucu

## 5. Veri Tabanı ve Veri Ambarı Karşılaştırması

ÇİHİ kapsamında ilişkisel veri tabanında tutulan veriler ve bu veriler üzerinde çalışan raporlar ile ÇİAİ kapsamında veri ambarında ve çok boyutlu küplerde tutulan veriler ve bu veriler üzerinde çalışan raporların karşılaştırmasını yaparsam Çizelge 5.2 ile gösterebileceğimiz sonuç kümesini gözlemlemekteyim.

Karşılaştırma için aylık istatistik sonuçlarını üreten, 19 sorgu ve her bir sorgu için tarihsel kümeleme fonksiyonlarını içeren hesaplanmış değerleri gösteren birer sorgu ile toplam da 38 adet sorgu barındıran bir raporu ele alalım. Rapor kapsamında yer alan sorgular Çizelge 5.1 ile gösterilmiştir.

19 adet ana sorgu gerçekleştirilen yazılım kapsamında bir rapor bünyesinde toplanmıştır. Toplam çalışma süresi 10 dk. dır. Aynı raporu ÇİHİ sistemi bünyesinde veri tabanı üzerinde YSD ile hazırlanacak sorgularla elde etmek oldukça maliyetlidir. Bazı sorgular için oluşturulacak sonuç kümesi birden çok rapor ile üretilebilmektedir. Bazı sorguların çalıştırılarak sonuç üretmesi 2 günü bulabilmektedir. Ayrı ayrı elde edilen rapor sonuçlarının tek bir rapora dönüştürülmesi ve bu veriler üzerinden istatistiksel sonuçlar çıkarılabilmesi için de ayrıca bir iş gücü harcanması gerekmektedir. Bu çalışmanın raporların alındığı ortamdan başka bir ortamda (word, excel v.b) yapılması gerekmektedir

Çizelge 5.1 Raporu oluşturan sorgu içerikleri

| Sorgu No | İçerik  |
|----------|---|
| 1        | Uzmanlık dallarına göre yatışı yapılan hasta sayıları |
| 2        | Yatak türlerine göre toplam yatak sayıları            |
| 3        | Günlük yatan hasta sayıları                           |
| 4        | Ortalama günlük yatan hasta sayıları                  |
| 5        | Yatak işgal yüzdeleri                                 |
| 6        | Ortalama yatış süreleri                               |
| 7        | Doğum türlerine göre doğum sayıları                   |
| 8        | Uzmanlık dallarına göre muayene sayıları              |
| 9        | Muayenenin yapıldığı bölüme göre muayene sayıları     |

Çizelge 5.1 Raporu oluşturan sorgu içerikleri (devam)

| <b>Sorgu No</b> | <b>İçerik</b>   |
|-----------------|---|
| 10              | Ameliyat türlerine göre ameliyat sayıları                   |
| 11              | Anestezi yapılarak gerçekleşen ameliyat sayıları            |
| 12              | EI Cerrahisi ameliyat sayıları                              |
| 13              | Patoloji türlerine göre gerçekleşen patoloji işlem sayıları |
| 14              | Laboratuvar gruplarına göre laboratuvar işlemleri sayısı    |
| 15              | Radyoloji tetkik sayıları                                   |
| 16              | İşlem türlerine göre acilde gerçekleşen işlem sayıları      |
| 17              | İşlem türlerine göre gerçekleşen tıbbi işlem sayıları       |
| 18              | Hazırlanan reçete sayıları                                  |
| 19              | Nedenlere göre gruplanmış ölüm sayıları                     |

Çizelge 5.2 Veri Tabanı ve Veri Ambarı Karşılaştırması

|                                    | <b>Veri Tabanı</b> | <b>Veri Ambarı</b> |
|------------------------------------|--------------------|--------------------|
| Toplam tablo sayısı                | 60                 | 31                 |
| Toplam Veri Sığası                 | 60 GB              | 30 MB              |
| Toplam Rapor Sayısı                | 30                 | 1                  |
| Raporların Ortalama Çalışma Süresi | 2 saat             | 10 dk.             |

## 6. SONUÇ

'Veri Ambarı' yaklaşımı ile, birden çok farklı veri kaynaklarında tutulan verilerin temizlenip, yeniden düzenlenerek ortak bir çatı altında toplanmasıyla oluşan veri kümeleri üzerinden hizmet sunulması günümüzün yaygın yaklaşımlarından birisidir. Veri ambarında tutulan ilişkisel tablolardan çok boyutlu küplere geçiş sağlanarak ÇİAİ teknolojileri ile verinin çok daha kullanışlı olması, verinin anlamlılığının artırılması ve veriye hızlı erişim sağlanmaktadır. Ayrıca veri madenciliği modelleri geliştirilerek karar destek kapsamında yeni bir bakış açısı sunulabilmektedir. Veri madenciliği algoritmaları, veriler arasındaki ilişkilerin ve kümelerin ortaya çıkarılmasını sağlayarak karar destek kapsamında yeni sonuç kümelerinin üretilmesini mümkün kılmaktadır.

Tez kapsamında, veri ambarı, ÇİAİ ve veri madenciliği yaklaşımları kullanılarak karar destek amaçlı raporlama aracı gerçekleştirilmiştir.

Sistemin geliştirilmesi esnasında;

- ÇİHİ sisteminde yer alan kaynak verilere erişilerek, verilerin temizlenmesi, eksik verilerin ayıklanması, verilerin dönüştürümü yapılarak veri ambarı üzerinde temiz verilerin oluşturulması,
- Hazırlanan veri ambarı ilişkisel tabloları kullanılarak ÇİAİ kapsamında çok boyutlu küplerin tasarlanması ve gerçekleştirimi,
- Küpler üzerinde tutulan verileri sorgulamak için ÇBİ cümleciklerinin hazırlanması,
- Hem veri ambarındaki ilişkisel tablolar üzerinde hem de çok boyutlu küpler üzerinde sınıflandırma veri madenciliği algoritmasının uygulanması,
- Veri ambarına, çok boyutlu küplere ve madencilik sonuç kümelerine erişimi sağlayarak sorgulama yapılabilmesini sağlayan yazılımın gerçekleştirilmesi,
- Gerçekleştirilen yazılım kapsamında oluşturulan ÇBİ cümleciklerine birden çok sorgunun yan yana ya da alt alta sonuç kümesi oluşturacak şekilde yapılandırılmasını sağlayan yeni komut yaklaşımlarının eklenmesi,

- ÇBİ cümlecikleri ile üretilen sorgu sonuçlarının hem grid bileşeni üzerinde hem de grafik çıktısı olarak gösterilmesinin sağlanması,
- ÇİAİ küplerine erişim için yaygın olarak kullanılan “Küp Editörü ve Pivot Table Servisi” araçlarına geliştirilen yazılım tarafından da ulaşılmasının sağlanması,
- Birden çok sorgunun yer aldığı raporların hazırlanarak Excel ürünü ile raporlanması,
- Veri madenciliği karar ağaçlarının hem ağaç yapısında görsel sunumu hem de grid bileşeni üzerinde gösteriminin sağlanması,

tasarlanmış ve başarıyla gerçekleştirimi yapılmıştır.

Geliştirilen sistemde kullanılan veriler, ‘Hastane Bilgi Yönetim Sistemine’ aittir. Fakat konfigürasyon dosyasında yapılacak değişikliklerle istenilen veri ambarına ve çok boyutlu küplere erişim sağlanılarak istenilen konu başlığı altında tutulan sistemlere bağlantı sağlanabilmektedir. Kod içerisinde gömülü konfigürasyon bilgileri olmadığı için sistem uygun bağlantı cümlecikleri ile farklı veri ambarları ile çalışabilmektedir. Bu sayede sistemin esnekliği sağlanmıştır.

## **KAYNAKLAR**

### Kitaplar:

- Agosta L., 1999, The Essential Guide to Data Warehousing, Prentice Hall
- Anahory, S., Murray, D. 1997. Data Warehousing in the Real World: A practical guide for building decision support systems. Addison-Wesley Longman, Inc.
- Barquin, R., Edelstein, H. A. (ed.). 1997. Building, Using, and Managing the Data Warehouse. Prentice Hall, New Jersey.
- Berson A., Smith S., 1997, Data Warehousing, Data mining, &OLAP , McGraw-Hill
- Bischoff J.,Alexander T.,1997, Data Warehouse, Practical Advice From The Experts, Prentice Hall
- Bishop C., 1996, Neural Networks for Pattern Recognition, Oxford Univ Press.
- Chen, Z., 2001, Data Mining and Uncertain Reasoning : An Integrated Approach, Wiley Inter-Science Publication
- Erik T.,1999, Microsoft OLAP Solutions
- Devlin, D. 1997. Data Warehouse from Architecture to Implementation. Addison-Wesley Longman, Inc.
- Han, Jiawei. 2001. Data mining : concepts and techniques. Morgan Kaufmann Publishers,San Francisco.
- Jarke M., Lenzerini M., Vassiliou Y., Vassiliadis P., 2000, Fundamentals of Data Warehouses
- Kachur, Richard J. 2000. Data warehouse management handbook. Prentice Hall, Paramus, NJ.
- Kimball R., 2002, The Datawarehouse Toolkit : The Complete Guide to Dimentional Modeling,
- Korth H.F., Silberschatz A., 1991, Database System Concepts, McGraw-Hill
- Mitchell T., 1997, Machine Learning, McGraw-Hill
- Perkovich, Dave, 2001, MCDBA SQL Server 2000, McGraw-Hill
- Seidman, C., 2001, Data Mining with Microsoft SQL Server 2000, Microsoft Press
- Spofford G.,2001,MDX Solutions With Microsoft SQL Server Analysis Services

Weiss S.M., Indurkha N., 1998, Predictive Data Mining: A Practical Guide, Morgan Kaufmann

Bilimsel periyodikler:

Ahmad I., Azhar S., Lukauskis P., 2004, Development of a decision support system using data warehousing to assist builders/developers in site selection, Decision Support Systems 37 (2004) 1 – 21

Berndt D.J, Fisher J.W., Hevner A.R., Studnicki J, Healthcare Data Warehousing and Quality Assurance, University of South Florida

Berndt D., Hevner A.R., Studnicki J., 2003, The Catch data warehouse: support for community health care decision-making, Decision Support Systems 35 (2003) 367– 384

Breault J.L., Goodall C.R., Fos P.J., 2002, Data mining a diabetic data warehouse, Artificial Intelligence in Medicine 26 (2002) 37–54

Body M., Miquel M., Bédard Y., Tchounikine A., 2002, A Multidimensional and Multiversion Structure for OLAP Applications, ACM 1-58113-590-4/02/001

Chau K.W., Ying Cao, Anson M., Zhang J., 2002, Application of data warehouse and Decision Support System in construction management, Automation in Construction 12 (2002) 213– 224

Chaudhuri S., Dayal U., 1996. An Overview of Datawarehousing and OLAP Technology

Chittaro L., Combi C., Trapasso G., 2003, Data mining on temporal data: a visual approach and its clinical application to hemodialysis, Journal of Visual Languages and Computing 14 (2003) 591–620

Edward F. Ewen, Carl E. Medsker, Laura E. Dusterhoft, 1999, Data Warehousing in an Integrated Health System; Building the Business Case , ACM 1999 I-581 13-120

Hwanga H., Kua C., Yen D., Cheng C., 2004, Critical factors influencing the adoption of data warehouse technology: a study of the banking industry in Taiwan, Decision Support Systems 37 (2004) 1 – 21

Kerkri E.M., Quantin C., Allaert F.A., Cottin Y., Charve P., Jouanot F., Y´etongnon K., 2001, An Approach for Integrating Heterogeneous Information Sources in a Medical Data Warehouse, Journal of Medical Systems, Vol. 25, No. 3, 2001

Knudsen L., 2005, Global gene mining and the pharmaceutical industry, Toxicology and Applied Pharmacology 207 (2005) S679 – S683

- Kusiak A, Caldarone C., Kelleher M., Lamb F., Persoon T., Burns A., 2004, Hypoplastic left heart syndrome: knowledge discovery with a data mining approach, *Computers in Biology and Medicine* 36 (2006) 21–40
- Lehmann H.P., Shortliffe E.H., 2003, Information Technology Support of Clinical Research: An Introduction, *Information Systems Frontiers* 5:4, 415–419, 2003
- Levene M., Loizou G., 2003, Why is the snowflake schema a good data warehouse design?, *Information Systems* 28 (2003) 225–240
- Michael O. Akinde M.O., Böhlen M.H., Johnson T., Lakshmanan L., Srivastava D., 2003, Efficient OLAP query processing in distributed data warehouses, *Information Systems* 28 (2003) 111–135
- Miller G.C., Dolan T.C., Crowson N., Stout S., 2002, Data warehousing and information management strategies in the clinical immunology laboratory, *Clinical and Applied Immunology Reviews* 3 (2002) 73–86
- Moshe Zviran M., Armoni A., Glezer C., 1998, HIS/BUI: A Conceptual Model for Bottom-Up Integration of Hospital Information Systems, *Journal of Medical Systems*, Vol. 22, No. 3, 1998
- Mullins I., Siadaty M., Lyman J., Scully K., Garrett C., Miller W., Muller R., Barry Robson B., Apte C., Weiss S., Rigoutsos I., Platt D., Cohen S., Knaus W., 2005, Data mining and clinical data repositories: Insights from a 667,000 patient data set, *Computers in Biology and Medicine*
- Niemi T., Nummenmaa J., Than P., 2001, Constructing OLAP Cubes Based on Queries, *ACM I-581 I3-437-1/01/0011*.
- Park C., Kim H., Lee Y., 2002, Finding an efficient rewriting of OLAP queries using materialized views in data warehouses, *Decision Support Systems* 32 (2002) 379–399
- Pedersen D., Riis K., Pedersen T.B., A powerful and SQL compatible data model and query language for OLAP, Department of Computer Science, Aalborg University
- Schubart J.R., Einbinder J.S., 2000, Evaluation of a data warehouse in an academic health sciences center, *International Journal of Medical Informatics* 60 (2000) 319–333
- Semenova T., 2004, Discovering patterns of medical practice in large administrative health databases, *Data & Knowledge Engineering* 51 (2004) 149–160
- Shoshani A., OLAP and Statistical Databases: Similarities and Differences



Vaisman A., Mendelzon A., Ruaro W., Cymerman S., 2004, Supporting dimension updates in an OLAPserver, Information Systems 29 (2004) 165–185

İnternet (www/ftp) kaynakları:

<http://www.olapreport.com/>

[only4gurus.com/v3/downloadcount.asp?id=6966](http://only4gurus.com/v3/downloadcount.asp?id=6966)

[http://www.thinkrealhard.com/data\\_analysis/books/pdfs/data\\_mining/SQL%20Data%20Mining.pdf](http://www.thinkrealhard.com/data_analysis/books/pdfs/data_mining/SQL%20Data%20Mining.pdf)

<http://www.angoss.com/>

Kullanıcı El Kitapları:

Oracle Database Documentation Library ,2005, Data Mining Concepts

Business Objects Query and Analysis User Manual

Oracle 9i Warehouse Builder User Manual

## ÖZGEÇMİŞ

Adı Soyadı : Selda Düzgünoğlu

Doğum Yeri : Isparta

Doğum Yılı : 1976

Medeni Hali : Evli

Eğitim ve Akademik Durumu:

Lise 1990-1993 Meram Anadolu Lisesi , KONYA

Lisans 1993-1997 Hacettepe Üniversitesi, Bilgisayar Bilimleri Mühendisliği  
Bölümü, ANKARA

Yabancı Dil:

İngilizce : Anlama ve okuma çok iyi, konuşma iyi

İş Tecrübesi :

1997-2000 : STM AŞ, Yazılım ve Tasarım Uzmanı

2000-2002 : SAMPAŞ AŞ, Sistem Çözümleyici ve Tasarım Uzmanı

2003-2004 : KMO AŞ, Yazılım Mühendisi

2004 -2006 : Hacettepe Üniversitesi Bilgisayar Mühendisliği Bölümü, Araştırma  
Görevlisi

