

T.C.
BEYKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ BİLİM DALI

KURUMSAL İŞ ZEKASI
(Yüksek Lisans Tezi)

Tezi Hazırlayan: **Yunus Nadi ERDEMİR**

T.C.
BEYKENT ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ BİLİM DALI

KURUMSAL İŞ ZEKASI
(Yüksek Lisans Tezi)

Tezi Hazırlayan:
Yunus Nadi ERDEMİR
Öğrenci No:
050820014

Danışman:
Dr. Rifat ÇÖLKESEN

YEMİN METNİ

Sunduđum Yüksek Lisans Tezimi, Akademik Etik İlkelerine bađlı kalarak, hi kimseden akademik ilkelere aykırı bir yardım almaksızın bizzat kendimin hazırladıđına and ierim.

04/05/2009



Aday : Yunus Nadi ERDEMİR

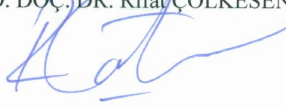
T.C.
BEYKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ
TEZLİ YÜKSEK LİSANS TEZ SINAV TUTANAĞI

25.05.2009

Enstitümüz Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği Bilim Dalı yüksek lisans öğrencilerinden 050820014 numaralı **Yunus Nadi ERDEMİR**' e "Beykent Üniversitesi Lisansüstü Eğitim - Öğretim Yönetmeliği" nin ilgili maddesine göre hazırlayarak, Enstitümüze teslim ettiği "**KURUMSAL İŞ ZEKASI**" tezini, Yönetim Kurulumuzun 30.04.2009 tarih ve 2009/05 sayılı toplantısında seçilen ve Fakülte binasında toplanan jüri üyeleri huzurunda, ilgili yönetmeliğin (c) bendi gereğince aday tarafından savunulmuş ve sonuçta adayın tezi hakkında **oybirliği** ile **Kabul** kararı verilmiştir. (60 dk.)

İşbu tutanak, 4 nüsha olarak hazırlanmış ve Enstitü Müdürlüğü'ne sunulmak üzere tarafımızdan düzenlenmiştir.

DANIŞMAN
YRD. DOÇ. DR. Rifat ÇÖLKESEN



ÜYE
Prof. DR. Esat HAMZAOĞLU



ÜYE
Yrd. Doç. Dr. Gökhan SİLAHTAROĞLU



TEŐEKKÜR

Tezimin konusunun belirlenmesi ve hazırlanmasında bana yardımlarını esirgemeyen Sayın Dr. Rıfat ÇÖLKESEN'e, katkılarından dolayı Murat Çetin Sav'a, Utku Şahin'e ve destekleri için aileme teşekkür ederim.

Yunus Nadi ERDEMİR

KURUMSAL İŐ ZEKASI

Tezi Hazırlayan : Yunus Nadi ERDEMİR

Özet

Günümüzde kurumlar rakiplerine göre Pazar paylarını artırabilmek ve güçlü bir rekabet avantajı sağlayabilmek için müşterisini çok iyi tanıyabilmek, müşterisine ürün, hizmet ve hizmetlerin sunumu anlamında farklılaşmak zorundadırlar. Bu çalışmada iş zekasını tüm katmanlarıyla inceleyerek kurumsal iş zekası yaratmanın getireceđi kazanımlar ve kurulum aşamaları konusunda bilgiler vermeye çalışılmıştır. Bankacılık üzerine örneklemeler yapılarak kurumlara yol gösterilmiştir.

Anahtar Kelimeler : İş Zekası, Veri Ambarı, OLAP, Veri Madenciliđi, Modelleme, Kurumsal Performans Yönetimi ve Kurumsal Karneleme

CORPORATE BUSINESS INTELLIGENCE

Presented by: Yunus Nadi ERDEMİR

Abstract

In today's competitive markets, motivations such as increasing market share and gaining competitive advantage force corporations to understand their existing and potential customers more than ever before. The key success factors lie in how the companies differentiate themselves in product design, delivery and associated customer service. This study analyzes all the aspects of business intelligence and provides information on the benefits, as well as the necessary investment and implementation processes at the enterprise level. Applications from the banking sector are provided as examples which could also be used as institutional guidance.

Key Words: Business Intelligence, Data Warehouse, OLAP, Modeling, Budget Planning and Consolidation and Balanced Scorecard.

İÇİNDEKİLER

Yemin Metni	
Jüri Sayfası	
Teşekkür	
Türkçe Özet ve Anahtar Kelimeler	
İngilizce Özet ve Anahtar Kelimeler (Abstract)	
İçindekiler	i
Tablolar Listesi	vi
Şekiller Listesi	vii
Kısaltmalar Listesi	ix
Giriş	1

I. BÖLÜM İŞ ZEKASI

1.1. İş Zekâsı (Business Intelligence) Tanımları	3
1.2. İş Zekasını Neden Kullanırız?	4
1.3. İş Zekası Döngüsü	4
1.4. Pratikte iş zekası nedir?	5
1.5. İş Zekası Raporlama Tanımları	5
1.5.1. Kapsam	5
1.5.2. Raporların Kullanımı	6
1.5.3. Rapor ve Analiz Tipleri	6
1.6. İş Zekası ve Sağladıkları	9
1.7. Gelişim Süreci	11

II. BÖLÜM VERİ AMBARI

2.1. Veri Ambarı	12
2.2. Veri Ambarı nedir?	12
2.3. Veri ambarı Özellikleri	13
2.4. Veri Ambarı ve İş Zekası	14
2.5. Ambar Verisinin Değişimi - Changing Warehouse Data	14
2.6. Veri Ambarının Amaçları	15
2.7. Veri Ambarı Bileşenleri	15

2.7.1. Kaynak Sistem (Source System)	15
2.7.2. Data Staging Area	15
2.7.3. Operational Data Store	16
2.7.4. Presentation Server	16
2.7.5. Dimensional Model	16
2.7.6. Data Mart	16
2.7.7. OLAP (On-Line Analytic Processing)	16
2.7.8. Metadata	16
2.8. Veri Ambarı Çatısı (Data Warehouse Framework)	17
2.8.1. Veri Ambarı Mimarisi	17
2.8.1.1. Operasyonel ve Dış Veri Kaynakları Katmanı	17
2.8.1.2. Bilgi Erişim Katmanı	17
2.8.1.3. Veri Erişim Katmanı	18
2.8.1.4. Veri Kılavuzu (Metadata) Katmanı	18
2.8.1.5. İşlem Yönetim Katmanı.	18
2.8.1.6. Uygulama İleti Gönderme Katmanı	18
2.8.1.7. Veri Ambarı Katmanı	19
2.8.1.8. Veri Sunum Katmanı	19
2.8.2. Mimari Seçenekler	19
2.8.2.1. Kurumsal Genişlikte Veri Ambarı (Enterprisewide Data Warehouse)	19
2.8.2.2. Veri Pazarı (Datamart)	20
2.8.2.2.1. Bağımsız Veri Pazarı Mimarisi (Independent Data Mart Architecture)	20
2.8.2.2.2. Bağımlı Veri Pazarı Mimarisi (Interconnected Data Mart Architecture)	21
2.8.3. Veri Ambarı Geliştirme Yaklaşımları	22
2.8.3.1. Yukarıdan Aşağıya Yaklaşım	22
2.8.3.2. Aşağıdan Yukarıya Gerçekleştirme	22
2.8.3.3. Birleştirilmiş Yaklaşım	23
2.9. Verinin Biçimlendirilmesi	24
2.10. Verinin Yapılandırılması	24

2.10.1. Gerçek Zamanlı Veri	25
2.10.2. Türetilmiş Veri	25
2.10.3. Uzlaşmış Veri	25
2.11. Veri Ambarı için Veri Modelleme	26
2.11.1. Veri Modelleme Teknikleri	26
2.11.1.1. Varlık-Bağıntı Modelleme	26
2.11.1.2. Boyutlu Modelleme	27
2.11.1.2.1. Olay	27
2.11.1.2.2. Boyut	27
2.11.1.2.3. Ölçek	28
2.11.1.2.4. Küp	28
2.11.1.2.5. Çok Boyutlu Veri Modellerinde OLAP Operasyonları	32
2.11.1.2.5.1. Roll-Up ve Drill-Down	32
2.11.1.2.5.2. Slice ve Dice (Dilimleme ve Çevirme)	33
2.11.1.2.6. Çok Boyutlu Veritabanları için Şemalar	33
2.11.1.2.7. Kavram Hiyerarşilerinin Tanıtılması	36
2.12. Veri Ambarı Tasarım ve Geliştirme	36
2.12.1. Veri Ambarı Tasarım Aşamaları	36
2.12.1.1. Business Model Tanımlama	36
2.12.1.1.1. Stratejik Analiz	36
2.12.1.1.2. Business Modelin Yaratılması	37
2.12.1.1.2.1. İş ihtiyaçlarının tanımlanması	37
2.12.1.1.2.2. Boyutlar ve Ölçeklerin Tanımlanması	38
2.12.1.1.2.3. Duyarlılıkların(Granularity) Belirlenmesi	39
2.12.1.2. Mantıksal (Logical) Modelin Tanımlanması	39
2.12.1.3. Boyutlu(Dimensional) Modelin Tanımlanması	40
2.12.1.3.1. Olayların (Facts) Tanımlanması	40
2.12.1.3.2. Eklenebilirlik (Additivity)	40
2.12.1.3.3. Zaman Boyutu Eklenmesi	41
2.12.1.4. Fiziksel Modelin Tanımlanması	42
2.13. Kaynakların Tanımlanması	43

2.14. Veri Kalitesi	43
2.15. Veri Ambarı Üretim Süreci (ETL - Extract, Transform, Load)	44
2.15.1. Veri Yakalama (Capture)	44
2.15.2. Veri Dönüştürme (Transform)	45
2.15.3. Uygulamaya Koyma (Apply)	45
2.16. Veri Ambarı Proje Yönetim Aşamaları	46

III. BÖLÜM

VERİ MADENCİLİĞİ

3.1. Veri Madenciliği (Data Mining)	48
3.2. Veri Madenciliği İşlevi	49
3.3. Veri Madenciliğinin Karar Verici için Yararları	49
3.4. Veri Tabanlarında Bilgi Keşfi (VBK) Süreci (Knowledge Discovery in Databases)	50
3.5. Veri Madenciliğindeki Problemler	52
3.6. Veri Madenciliğinde Kullanılan Yöntemler	54
3.6.1. İstatistiksel Yöntemler	54
3.6.2. Bellek Tabanlı Teknikler	56
3.6.3. Genetik Algoritmalar	57
3.6.4. Yapay Sinir Ağları	57
3.6.5. Karar Ağaçları	57

IV. BÖLÜM

KURUMSAL PERFORMANS YÖNETİMİ

4.1. Kurumsal Performans Yönetimi	58
4.2. Anahtar Performans Göstergeleri (KPI) Kavramı	58
4.3. Balanced Scocard (Kurumsal Karne)	59

V. BÖLÜM

BANKACILIK SEKTÖRÜNDE ÖRNEK BİR İŞ ZEKASI UYGULAMASI

5.1. Veri Ambarı Tasarım Özellikleri	62
5.2. Veri Madenciliği Uygulamaları	70
5.2.1. Ürün Eğilim Modellemeleri	71
5.2.1.1. Çapraz Satış (Cross Sell)	71
5.2.1.2. Ek Satış (Up Sell)	72

5.2.1.3. Dizi Analizi (Sequence analysis)	72
5.2.1.4. Sepet Analizi (Market-Basket analysis)	72
5.2.2. Kümeleme	72
5.2.3. Terk Analizi (Churn)	72
5.2.4. Davranışsal Skorlama (Behavioral skoring)	73
5.2.5. EAD hesaplamaları	74
5.2.6. LGD hesaplamaları	74
5.2.7. CCF hesaplamaları	74
5.3. Analitik Modelleme ve Raporlama Uygulamaları	74
5.3.1. Segmentasyon	74
5.3.2. Hesap İşletim Ücreti	76
5.3.3. Gişe Satış Modülü	76
5.3.4. Müşteri Sadakat Uygulaması (Loyalty)	77
5.4. Analitik ve Statik Raporlar	78
5.5. Veri Ambarı Modelleme	80
5.6. BI Raporlama Platformu Modelleme	87
5.7. ETL Süreci	88
5.8. İş Zekası Projesi Yönetim Önerileri	89
5.9. İş Zekası Platform ve ETL Tool Seçimi	90
5.10. İş Zekası Raporlama ve Dashboard Örnekleri	106
SONUÇLAR	115
KAYNAKLAR	116
ÖZGEÇMİŞ	118

TABLULAR LİSTESİ

<u>Tablo No.</u>	<u>Sayfa</u>
1. İş süreçleri Matrisi Örneği	37
2. Veri Ambarı Proje Yönetim Adımları ve Detaylar	46

ŞEKİLLER LİSTESİ

<u>Şekil No.</u>	<u>Sayfa</u>
1. İş Zekası Raporlama	6
2. Ansal Raporlama Örneği – 1	8
3. Ansal Raporlama Örneği – 2	9
4. İş Zekası Araçları	10
5. Verinin İş Zekası süreçlerinden geçerek bilgiye dönüşüm süreci	10
6. Üç Boyutlu Küp	29
7. Küpün Bir Zarı (Fon)	30
8. Üç Boyutlu Bir Küp	31
9. Üç boyutlu küpten alınan bir zar(3D>2D)	31
10. OLAP verisini temsil eden veri küpü	31
11. Satış analizi için üç boyutlu bir verinin ilişkisel gösterimi	31
12. Roll-Up Örneği	32
13. Drill-Down örneği	33
14. Yıldız Şema	34
15. Kar Tanesi Şema	35
16. Boyutlar ve Ölçekler	38
17. ER Model	39
18. Çoklu Zaman Boyutu	42
19. Veri Ambarı Proje Yönetim Aşamaları	46
20. Bilgi keşfi sürecinde veri madenciliği (Data Mining in Knowledge Discovery)	54
21. Kurumsal Karne Basamakları	63
22. Veri Ambarı Modelleme Diyagramı (Party Base)	80
23. Veri Ambarı Modelleme Diyagramı (Adress Base)	81
24. Veri Ambarı Modelleme Diyagramı (Contract Base)	82
25. Veri Ambarı Modelleme Diyagramı (Contract Deposit)	83
26. Veri Ambarı Modelleme Diyagramı (Contract Loan)	84
27. Veri Ambarı Modelleme Diyagramı (Contract Investment)	85

28. Veri Ambarı Modelleme Örneđi (GL Account)	86
29. BI Raporlama Platformu Modelleme Örneđi	87
30. BI ETL Süreç Örneđi	88
31. Statik Raporlama Örnek-1	106
32. Statik Raporlama Örnek-2	107
33. Drill-Down Raporlama Örneđi - Segment	108
34. Drill-Down Raporlama Örneđi – Şube Detayı	108
35. Drill-Down Raporlama Örneđi – Müşteri Detayı	109
36. Analitik Raporlama Örnek-1	109
37. Analitik Raporlama Örnek-2	110
38. Analitik Raporlama Örnek-3	110
39. Dashboard Örnek-1	111
40. Dashboard Örnek-2	112
41. Dashboard Örnek-3	113
42. Dashboard Örnek-4	114

KISALTMALAR LİSTESİ

BI	Business Intelligence
BDDK	Banka Düzenleme Denetleme Kurulu
CRM.	Customer Relation Manager
DK	Defteri Kebir
DEK	Döviz Endeksli Krediler
ER	Entity – Relationship
ETL	Extract – Transform – Load
FTP	File Transfer Protocol
KPI	Key Performance Indicator
ODS	Operational Data Store
OLAP	Online Analytical Process
RFP	Request For Proposal
TCMB	Türkiye Cumhuriyeti Merkez Bankası
VA	Veri Ambarı
VB	Varlık – Bağını
Vb.	Ve Benzeri
VM	Veri Madenciliği
VTBK	Veri Tabanlarında Bilgi Keşfi

GİRİŞ

Rekabetin her geçen gün biraz daha arttığı günümüzde, müşterileri ile her noktada olan etkileşimlerini kendi menfaatlerine kullanabilen kuruluşlar güçlü bir rekabet avantajı kazanmaktalar.

Kurumların yönetim ve analiz kadrosu bugün birçok veriyle uğraşmak zorunda kalmaktadır. Bu noktada yapılması gereken şey kurumların içerisinde yer alan veriyi analiz etmek, onlarla birlikte bunu bilgiye dönüştürmek ve bu bilgi çerçevesinde de kurumların kendilerine değer katmasını sağlamak olmalıdır. Günümüzde, kurumlar için müşteri en önemli unsurlardan bir tanesidir. Yöneticilerin kendi vizyonlarını, stratejilerini çizerken gerçekten bilgiye dayanarak karar vermelerini istiyoruz.

Krizlerin çok yoğun yaşandığı ülkemizde krizden etkilenen çok sayıda kurum oluyor. Bunda kurumların zamanında kârlarını ölçememiş olmaları en büyük etken olmaktadır. Yöneticilerin önüne çok fazla veri gelmekte, gerçek anlamda analiz edilmiş bilgi olmadığı için süreç yanlış yönetilebilmektedir. Burada kurumların ve yöneticilerin ihtiyacı sadece anlık durumu değerlendirmek değil geleceğe dönük tahminleme yapabilmek olmalıdır. Kurumlar, beni hangi müşteriler bırakacak, ileriki dönemde hangi müşterilerim ödemesini yapabilir ya da benim ödemelerimi durdurabilir gibi sorulara hemen yanıt alabilmelidir.

Bugün baktığımızda, pazardaki en kritik şey kurumların iş yapma biçimlerinin değişmesidir. Eskiden kurumlar ürün merkezli işler geliştirirken, bugün iş süreçlerini müşteri merkezli olarak geliştirmektedir. Bütün kurumların %100 müşteri memnuniyeti sağlamak gibi bir misyonu olması mümkün değildir. Amaç kârlılık ve kârlı müşteriyi kaybetmemektir. Yani iki noktada iş süreci değişimi yaşanıyor. Birinci aşama gerçek anlamda müşteri kârlılığı, ikinci aşama ise kârlı müşteriyi elde tutmak. Tabii bu değişim kurumların elde ettiği yatırım maddelerinin ölçülmesini farklılaştırıyor.

Pazarlama otomasyonunun tanımı iş süreçleriyle teknolojinin evlendirilmesidir. Bu sistemin amacı bilgiye ulaşımı arttırmak, iletişimdeki istikrarı sağlamak ve müşteriye doğru kanallardan ulaşabilmektir. Bu pazarlama yönetiminin önünde bulunan büyük bir zorluk. Bu evlendirme yapılırken veri yönetimini kullanmak en kritik unsurdur. Ayrıca kurumlar tek bir müşteri görüntüsü sağlamalıdır. Tek bir müşteri görüntüsü yaratabilmek için bir çok kanal üzerinden gelen ve giden veriler tek bir noktada toparlanmalı, doğru analitik araçlarla analiz edilmeli ve gerçek anlamda tahminsel modellemeler çıkarılmalıdır. Bu bağlamda müşterilerin düşüncelerini ve tutumlarını anlamak, ne söylediklerini, ne yaptıklarını da bilmek çok büyük önem arz etmektedir. Operasyonel sistemdeki geçmiş veri ve demografik verinin bileşiminden

oluşan veri ile müşteri ile ilişkide bulunulan pek çok kanaldan (çağrı merkezi, internet, ATM, şube, vb.) gelen ilişkisel verileri kullanarak hem müşteri davranışlarını anlama hem de müşteri davranışını tahmin etme olanağı veren İş Zekası çözümlerini kullanarak karar verme süreçlerimizi desteklemek günümüz koşullarının en önemli sürecidir.

Karar destek sistemleri farklı amaçlara hizmet eden birçok farklı uygulamanın genel adı olarak kullanılmaktadır. Bunlar raporlama araçları, veri ambarı, veri madenciliği, OLAP, Kurumsal Karne, CRM gibi uygulamalardır. Eskiden eğer bu uygulamaların hepsini kullanmak istiyorsak çok yapılı altyapı mimarisi oluşturmak ve entegrasyon sağlamak zorunluluk gerektirmekteydi.

Günümüzde, iş zekası adını verdiğimiz ve tüm bu uygulamaları (Analitik uygulamalar, Bütçe, Karlılık, Risk Yönetimi, Veri Ambarı, OLAP, Veri Madenciliği, Kurumsal karne, ETL ve CRM) içeren ve uçtan uca tek bir platform üzerinden sunabilen teknoloji kullanılmaktadır. Bu özelliği ile iş zekası platformları çok parçalı altyapı mimarisini azaltmak suretiyle ciddi maliyet, entegrasyon ve yönetim tasarrufları sağlamaktadır.

En önemlisi rekabetçi iş piyasalarında kurumlara karar alma, strateji belirleme, bütçe planlama, ölçüleme ve yeniden değerlendirme konularında büyük avantajlar sağlamaktadır.

Tezimde iş zekası araçları hakkında genel bilgiler verdikten sonra, bu araçlara ait bankacılık sektörü için hazırlanmış örneklemeler ve iş zekası platformu ve ETL tool seçimi konusunda sorulması gereken soru setleri ve BI proje yönetim önerileri konularından bahsedeceğim. Bu konu kurumlar için çok önemlidir. Yanlış seçimler bu tür projelerin başarısız olmasına neden olur. Amacım iş zekası araçlarını tanıtmak, kurumların iş zekası platformu ve ETL tool seçiminde doğru seçim yapabilmelerine yardımcı olabilmektir. Sonrasında iş zekası projesi safhalarından, kurulum aşamalarından bahsetmek ve yol göstermektir.

İlk 4 bölümde iş zekası araçları hakkında teorik ve bankacılık sektöründe uygulamaya dönük bilgiler aktaracağım. 5. Bölümde iş zekası araçları ile bankacılık sektörü için hazırlanmış olduğum örneklemeleri, platform/ETL tool seçimi ve soru kümeleri, proje ve kurulum aşamaları konularında bilgilendirme yapmaya çalışacağım.

I. BÖLÜM : İŞ ZEKASI

1.1. İş Zekâsı (Business Intelligence) Tanımları

İngilizce'de Intelligence kelimesinin iki anlamı bulunmaktadır. Birinci anlama göre Intelligence, öğrenme ve anlama yeteneğini ifade eden zekadır. Diğer ise, enformasyon toplama anlamında kullanılmakta ve dilimizde özellikle istihbarat ile ifade edilmektedir.

BI, enformasyona bilgi teknolojileri destekli her türlü erişimi ve karar destek amacıyla enformasyonun analizini gerçekleştiren tüm süreçleri ifade etmektedir.(Gartner Group)

İş zekâsı, dağıtık veriler kullanılarak stratejik karar alma durumunda olan kişilere bilgi üretilmesidir. Veri analizi bakış açısı ile kişilere kararlarını belirlemek ve varsayımlarda bulunmalarına yardımcı olabilmek amacıyla anlamlı bilgilerin toplanması sürecidir. Bilgi bakış açısı ile verinin incelenmesi, keşfedilmesi ve dönüştürülmesi ile bilginin elde edilmesidir.

Temel olarak iş zekâsı, kurumların karar verme ve yönetim kabiliyetlerinin artırılmasına yardımcı olan, çok sayıda verinin kurumsal bilgiye dönüştürülmesini ve böylelikle kurumların rekabet ortamında avantaj sağlamasına yönelik kavramlar, metotlar, süreçler ve yazılımlar bütünüdür.

Özetlersek,

BI, kurumların karar verme süreçlerini etkileyen ve optimize eden

- Tüm araçların kullanımını,
- Verilerin toplanmasını,
- Saklanması,
- Düzenlenmesini,
- Analiz edilmesini,
- Kolay erişilmesini,
- Görselleştirilmesini,
- Planlama yapılmasını,
- Strateji belirlenmesini ve ölçülmesini,
- Verilerin en etkin ve kolay biçimde yönetilmesini ve
- Kritik yönetim kararlarının verilmesini sağlayacak

tüm süreçleri kapsayan bir uygulama ve teknolojiler bütünüdür.

BI veriyi bilgiye dönüştüren teknoloji ve süreçler topluluğudur.

1.2. İş Zekasını Neden Kullanırız?

İş Zekasını kullanma nedenleri aşağıdaki başlıklar altında toplanmıştır:

- o İşimizi daha performanslı yapabilmek için,
- o Daha karlı bir firma olmamız için,
- o Müşterilerimize daha iyi hizmet vermemiz için,
- o Müşteri ve ürünlerimizi gruplayarak ve bu grupları birbiriyle ilişkilendirerek daha etkin satış ve pazarlama yapmamız için,
- o Verimliliğimizi artırmak için,
- o Hangi kanaldan, hangi müşterilere, ne şekilde hizmet vermek, hangi ürünleri pazarlamak gibi soruları cevaplamak için,

Kısaca,

- o İşimizi anlamak, işimizle ilgili bilgi sahibi olmak
- o Doğru kararlar vererek işimizi doğru yönetmek için,

İş Zekası sistemlerini kullanmak isteriz

1.3. İş Zekası Döngüsü

İş Zekası döngüsü üç başlık altında incelenir. Bunlar :

- o Veriden Enformasyona: Veriler çeşitli operasyonel ortamlardan alınarak entegre edilmekte, ve bir veri ambarında depolanmaktadır. Bu veri seçme, alma ve entegre etme süreci, veriyi yeni bir ürüne dönüştürmektedir (Enformasyon).
- o Enformasyondan Bilgiye: Kullanıcılar, çeşitli raporlama ve analiz araçları (sorgu ve rapor araçları, OLAP araçları, dashboard'lar, veri madenciliği araçları) sayesinde veri ambarındaki enformasyona istedikleri zaman erişebilmektedirler. Bu yapılan analizler ve araştırmalar sonucunda enformasyon bilgiye dönüşmektedir.
- o Bilgiden Kurallara ve Aksiyonlara: Bu anlayış ve kavrayışlar kullanıcıların bilgiyi kurallara ve aksiyonlara dönüştürmesini sağlamaktadır. Bu kurallar basit (envanter 25 birimin altına düştüğünde 50 birim daha ısmarla) olabildiği gibi, skorlama modelleri, segmentasyon veya çapraz satış modelleri gibi komplike de olabilir.

1.4. Pratikte iş zekası nedir?

İş zekasının genel, iş aktiviteleri, kurumsal performans ve resmi kurallar açısından içeriği aşağıda verilmiştir :

- o Genel Olarak
 - Statik ve Parametrik Raporlama
 - Ansal Sorgulama Ortamları
 - OLAP Analizleri
 - Veri Madenciliği
- o İş Aktiviteleri İzleme
 - Dashboard'lar
 - Alert'ler, Alarmlar
- o Kurumsal Performans Yönetimi
 - KPI Metrikleri İzleme
 - Balanced Scorecard
- o Uyum Sağlanması Gereken (Resmi) Kurallar
 - BDDK
 - Basel II

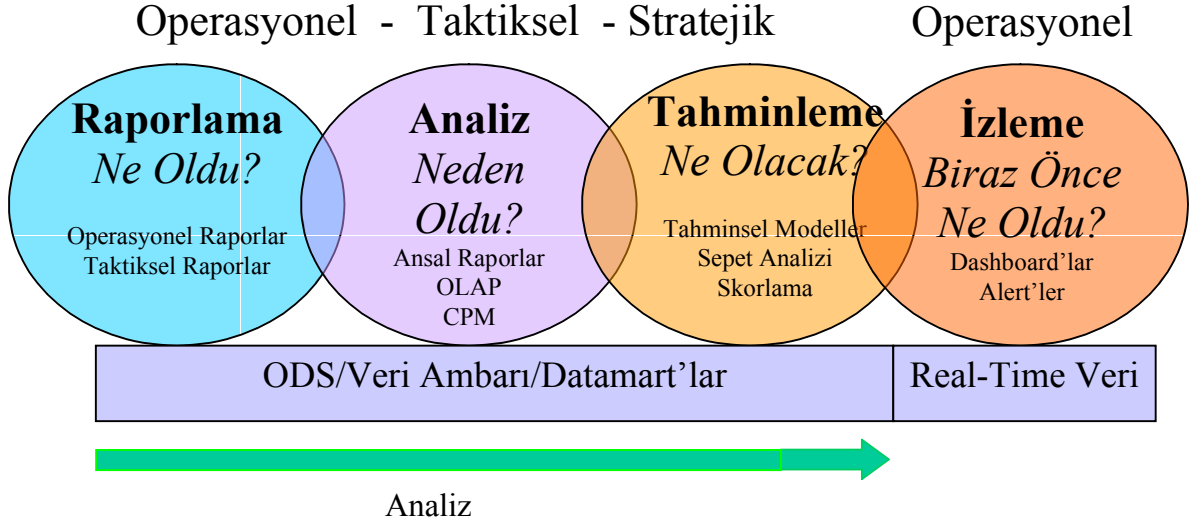
1.5. İş Zekası Raporlama Tanımları

1.5.1. Kapsam

İş zekası raporlama kapsamı amacına göre operasyonel, taktiksel ve stratejik açıdan gruplanabilir.

- o Operasyonel Raporlama
 - Tek sistem bakış açısını içerir.
 - En iyi ODS'ten sağlanabilir.
 - Amaçlanan, operasyonel sistemin desteklenerek en iyi şekilde çalışmasının sağlanmasıdır.
- o Taktiksel Raporlama
 - Entegre bakış açısını içerir.
 - En iyi Veri Ambarından sağlanabilir.
 - Amaçlanan, KPI'ların monitor edilmesidir.
 - Kısa vadeli kararlar için kullanılır.
- o Stratejik Raporlama
 - Entegre bakış açısını içerir.
 - En iyi Veri Ambarından sağlanabilir.

- Amaçlanan, verilerin uzun vadedeki gidişatının ölçülmesidir.
- Uzun vadeli kararlar için kullanılır.



Şekil 1.1 İş Zekası Raporlama

1.5.2. Raporların Kullanımı

Raporların kullanıcı bazlı kullanımı için belirlenmiş özellikler aşağıda anlatılmıştır :

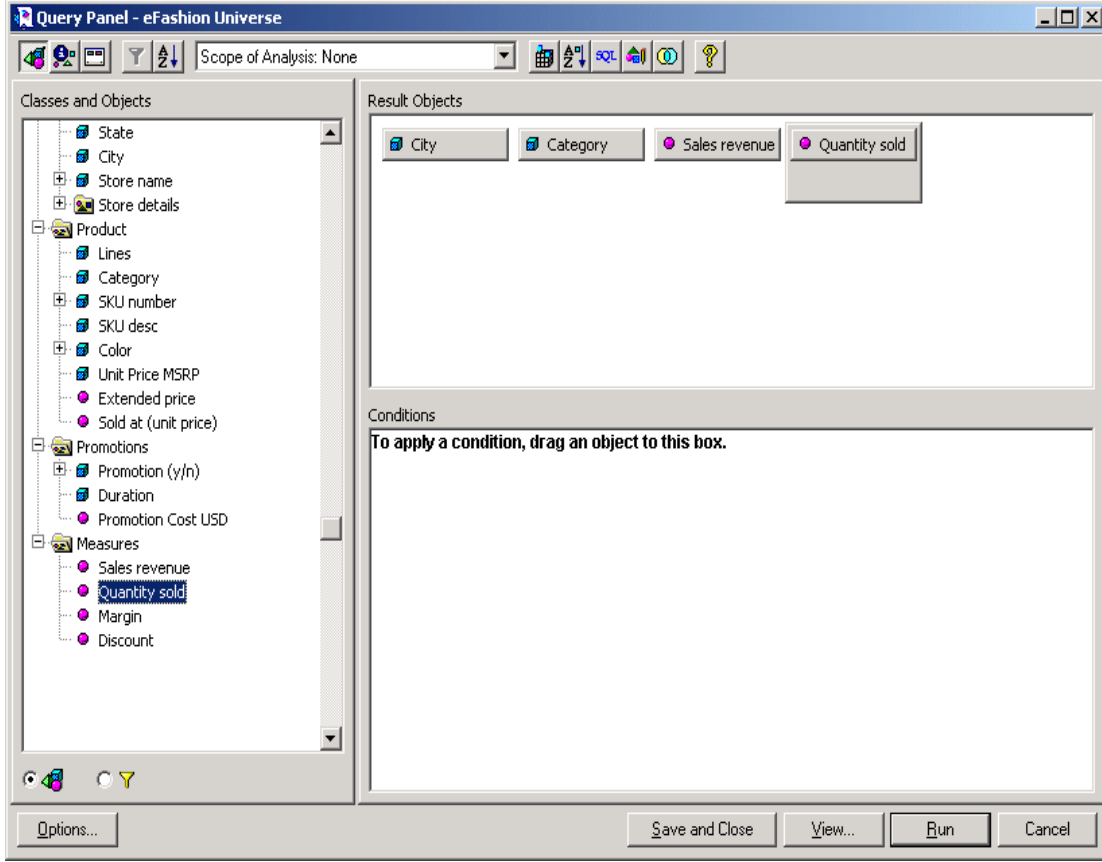
- Kullanıcıların çoğunluğu için:
 - İşlerinin sadece %10'u raporlamadır.
 - Genellikle statik raporlar ve hazır OLAP view'lerini kullanırlar.
- Bilgi çalışanları için (ürün yöneticileri, kredi ve risk yöneticileri) :
 - İşlerinin %30'u raporlama ve analiz aktivitelerini kapsar.
 - Genellikle parametrik raporları ve bazen OLAP küplerini kullanırlar.
- Analistler için:
 - İşlerinin %80'ine varan bir oranda raporlama ve analiz ile uğraşırlar.
 - Genelde ansal raporları ve OLAP küplerini kullanırlar.

1.5.3. Rapor ve Analiz Tipleri

Rapor ve analiz tipleri içerikleri ile aşağıda detaylandırılmıştır :

- Basit OLAP:
 - Kullanıcının profiline göre yetkili olduğu önceden hazırlanmış OLAP görünümleri liste şeklinde kendisine sunulur.
 - Sadece çok kısıtlı sayıda drill-down veya slice-and-dice alternatifleri mevcuttur.

- Bu görünümle parametrik raporlar şeklinde de kullanıcıya sunulabilir.
- Kontrollü Ansal Raporlar:
 - Kullanıcı genelde tek sistemden alınmış olan kısıtlı sayıda sahadan istediği raporu oluşturabilir. İleri derecede formatlama yapamaz.
- OLAP Analizi:
 - Kullanıcı profiline göre yetkili olduğu OLAP küplerine direkt olarak erişebilir ve kendi view'larını kendi oluşturabilir.
 - Her türlü drill-down, slice-and-dice ve drill-through olanaklarından faydalanabilir.
- Ansal Raporlar:
 - Kullanıcı birçok değişik sistemden alınmış verileri içeren çok sayıdaki sahadan istediği raporu oluşturabilir. İleri derecede formatlama yapabilir.
- Veri Madenciliği:
 - Kullanıcı birçok değişik sistemden alınmış verileri içeren çok sayıdaki sahaya erişerek bunları Data Mining ortamına transfer edebilir.
 - Bu verilere çeşitli transformasyonlar uygulayarak yeni değişkenler oluşturabilir.
 - Bu değişkenleri kullanarak çeşitli algoritmaları çalıştırabilir ve yeni Mining modelleri oluşturabilir.



Şekil 1.2 Ansal Raporlama Örneği – 1

Yukarıdaki ansal rapor örneğini çalıştırdığımızda aşağıdaki Şekil 1.3 de görülen raporu elde ederiz.

City	Category	Sales revenue	Quantity sold
Austin	2 Pocket shirts	\$67.824	361
Austin	Belts,bags,wallets	\$140.300	972
Austin	Bermudas	\$9.561	59
Austin	Boatwear	\$21.459	128
Austin	Cardigan	\$63.248	484
Austin	Casual dresses	\$42.180	296
Austin	Day wear	\$121.707	919
Austin	Dry wear	\$17.190	81
Austin	Evening wear	\$123.440	862
Austin	Fancy fabric	\$10.678	64
Austin	Full length	\$15.443	81
Austin	Hair accessories	\$26.350	173
Austin	Hats,gloves,scarw	\$134.757	1.022
Austin	Jeans	\$22.896	124
Austin	Jewelry	\$399.359	2.631
Austin	Long lounge pant	\$16.386	128
Austin	Long sleeve	\$92.185	590
Austin	Lounge wear	\$47.140	309
Austin	Mini city	\$11.081	56
Austin	Night wear	\$3.143	15
Austin	Outdoor	\$24.679	147
Austin	Pants	\$4.634	24
Austin	Party pants	\$22.066	118
Austin	Samples	\$71.108	536
Austin	Shirts	\$5.170	28
Austin	Short sleeve	\$114.156	637

Şekil 1.3 Ansal Raporlama Örneği – 2

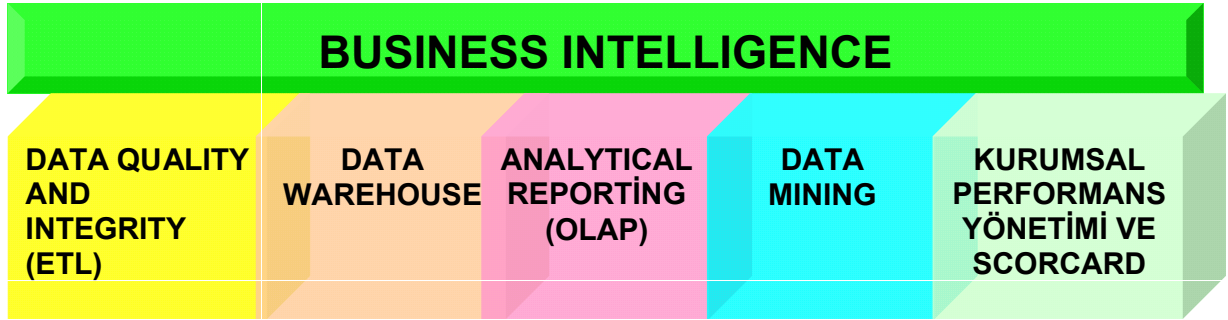
1.6. İş Zekası ve Sağladıkları

İş zekâsı, Müşteri İlişkileri Yönetiminde (Customer Relationship Management - CRM), müşteri karlılık değerlerinin ve bu değer oluşum sebeplerinin belirlenmesine ve böylelikle kurum açısından belirli bir eşik değeri üstünde değer ifade eden müşterilerden elde edilen faydanın artırılmasına yönelik çalışmalara destek verecek araçları içermektedir. Bu araçlar, daha önce de belirtildiği üzere kullanıcı gruplarının genişlemesine örnek olarak, kurumun arka planında bulunanlar, ön planda müşteri ile etkileşimde bulunan kişilerin bir araya gelerek iş zekâsından hizmet almasını sağlamaktadır.

İş zekası platformu :

Kurum performansının ölçümlenebilmesi amacıyla kurum sistem verilerinin tüm ilişki ağlarıyla tarifiyle model çerçevesinde depolandığı veri ambarı altyapısı üzerinde çalışacak ve kurumun işlem performansının finans, müşteri ve çalışan perspektifleri ile günlük ve tarihsel boyutlarda incelenebilmesine imkan sağlayarak, yönetimin karar alma sürecini destekleyecektir.

Hedef, kurumların sahip oldukları çalışma biçimini iş zekası yöntemleriyle değiştirmektir. Aşağıdaki şekil 1.4. İş Zekası araçlarını göstermektedir.

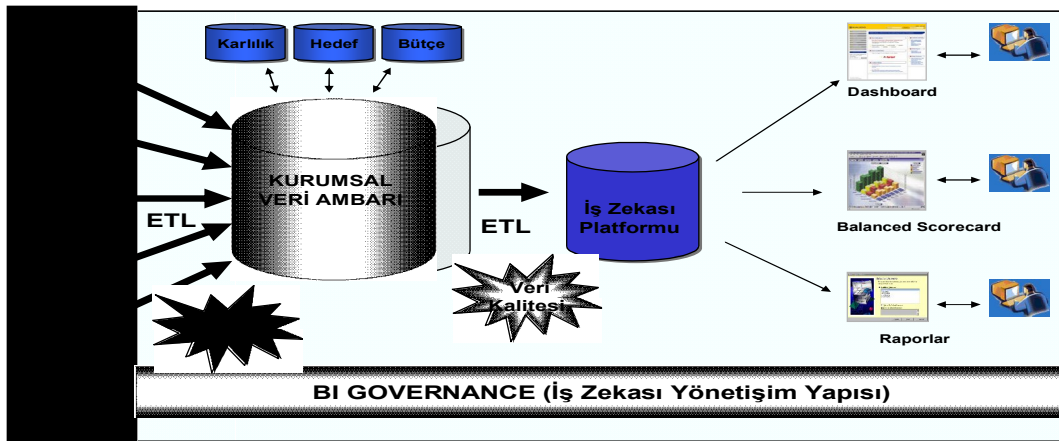


Şekil 1.4. İş Zekası Araçları

İş Zekası ile ulaşılanlar :

- Taktiksel kullanımdan →stratejik bilgi kullanımına geçiş,
- Parçalanmış veriler yerine →standartize edilmiş veriler ile analiz imkanı,
- Kişisel ve özel bilgiden →yaygın ve genel kullanımlı bilgiye geçiş,
- Veri yığından →bilgi kurumuna geçiş.

Verinin Operasyonel sistemlerden alınarak İş Zekası süreçlerinden geçip bilgiye dönüşüm süreci ve kullanıcıya sunumu aşağıdaki şekil 1.5 de gösterilmiştir. Bu süreç ileriki konularda detaylandırılmıştır.



Şekil 1.5. Verinin İş Zekası süreçlerinden geçerek bilgiye dönüşüm süreci

1.7. Gelişim Süreci

Gelişen dünyada rekabetin artması üzerine kurumlar doğru ve hızlı karar almak, uygulamak ve ölçümlemek için sistem arayışlarına yönelmiştir. Karar verme süreci, iş dünyasında etkili olan konuların incelenerek ileriye dönük üretim, pazarlama, bütçe, strateji, personel gibi birçok konuda öngöründe bulunmayı gerektirir. Karar, kurumun geçmiş işletme tecrübeleri üzerine, günümüzün verileri incelenerek verilmektedir. Karar verme sürecinin başarılı bir karara dönüşmesinde, geçmişteki tecrübelerin bulunması ve anlamlandırılması büyük önem taşımaktadır. Kurumların operasyonel verilerinin içerisinde gizli kalmış bilgilerin, alınacak karar sürecine faydalı hale getirilmesi, incelenmesi ve işlenmesi süreci Karar Destek Sistemleri tarafından ele alınmaktadır.

Karar destek sistemleri, basit yöntemler ile karmaşık algoritmaları bir arada kullanabilen sistemlerdir. Veri tabanından yapılan sorgular ile üretilen ilk aşama raporları, ilk akla gelen karar destek yöntemi olarak düşünülebilir. Bununla beraber, veri üzerinde matematiksel, istatistiksel ve sezgisel yöntemlerin kullanımını, karar destek sistemlerinin karar sürecinde, insan zekâsının kavrama yeteneğinin ötesine geçmesini sağlamaktadır.

Karar destek sistemleri farklı amaçlara hizmet eden birçok farklı uygulamanın genel adı olarak kullanılmaktadır. Bunlar raporlama araçları, veri ambarı, veri madenciliği, OLAP, Kurumsal Karne, CRM gibi uygulamalardır. Eskiden eğer bu uygulamaların hepsini kullanmak istiyorsak çok yapılı altyapı mimarisi oluşturmak ve entegrasyon sağlamak zorunluluk gerektirmekteydi.

Günümüzde, iş zekası adını verdiğimiz ve tüm bu uygulamaları (Analitik uygulamalar, Bütçe, Karlılık, Risk Yönetimi, Veri Ambarı, OLAP, Veri Madenciliği, Kurumsal karne, ETL ve CRM) içeren ve uçtan uca tek bir platform üzerinden sunulabilen teknoloji kullanılmaktadır.

Bu özelliği ile iş zekası platformları çok parçalı altyapı mimarisini azaltmak suretiyle ciddi maliyet, entegrasyon ve yönetim tasarrufları sağlamaktadır.

En önemlisi rekabetçi iş piyasalarında kurumlara karar alma, strateji belirleme, bütçe planlama, ölçümleme ve yeniden değerlendirme konularında büyük avantajlar sağlamaktadır.

II. BÖLÜM : VERİ AMBARI

2.1. Veri Ambarı

Veri ambarı oluşturulurken karar verilmesi gereken önemli bir konu verinin nasıl yapılandırılacağıdır. Verinin yapılandırılması, veri ambarcılığında en önemli konulardan biri olmasının yanı sıra, veri ambarını işlevsel veri tabanlarından (operational databases) ayıran noktalardan biridir.

Veri ambarı ilişkilendirilmiş veriler üzerinde sorgular ve analizler için kullanılabilir, bütünleşmiş bilgi deposudur. Standford Üniversitesine göre: Veri ambarı, başlangıçta farklı kaynaklardan gelen verinin üzerinde daha etkili ve daha kolay sorgu ve analiz yapılmasını sağlamaktadır.

Veri Ambarları, kurumların pazarlama bölümünden üretime, geleceğe dönük tahminler yapmada, sonuçlar çıkarmada ve kurumların yönetim stratejilerini belirlemede kullanılmakta olan bir sistemdir.

2.2. Veri Ambarı nedir?

Veri amarlama, kurumlara verilerini, sistematik bir şekilde organize etme, anlama, stratejik kararlarını vermede kullanmaları için mimariler ve araçlar sunmaktadır. Organizasyonlar, bugün global dünyanın rekabetçi, hızlı gelişen ortamında verimli, karlı, hızlı ve doğru karar verebilme yeteneklerini kazanmak ve kullanmak için veri ambarı sistemlerinin değerli araçlar olduğu kanısına varmıştır. Kurumlar, her alanda rekabetin oluşması ile birlikte, veri amarlama ve iş zekası ürünlerinin mutlaka sahip olunması gereken bir pazarlama silahı ve müşterilerin ihtiyaçlarını daha çok öğrenerek onları elde tutmanın bir yolu olduğu düşüncesinden hareket ederek son birkaç yılda bu konuda büyük yatırımlar yapmaktadırlar.

“Veri Ambarı nedir?” Sorusunun cevabı çok çeşitlilik göstermekle birlikte kesin olmayan biçimde söylenirse, veri ambarı sistemleri, çeşitli uygulama sistemlerinin bütünleşmesine imkan tanır. Bu sistemler de bilginin işlenmesini, analiz için pekiştirilmiş olan tarihi veriye ilişkin sağlam bir platform sağlayarak destekler.

Veri ambarları, geleneksel sorgulama ve raporlama yöntemlerinin yanısıra, Veri Madenciliği, çok boyutlu çözümleme (multi-dimensional analysis), kurumsal karneleme, bütçe, karlılık, CRM gibi güçlü veri çözümleme teknikleri için altyapı sağlar. Bu tekniklerin kullanılması sayesinde, karar verme sürecinde gerekli olan bilgiye daha kolay erişim sağlanabilmektedir.

Özetle, veri ambarı stratejik kararları verme konusunda bir kurumun ihtiyacı olan bilgiyi depolayan ve karar destek veri modelinin fiziksel bir sunumu gibi çalışan, anlamsal olarak tutarlı bir veri deposudur. Veri ambarı aynı zamanda, sıklıkla, yapısal ve/veya

planlanmamış sorgular, analitik raporlar ve karar vermeyi desteklemek için, çeşitli farklı türdeki kaynaklardan verinin bütünleştirilmesi ile oluşturulan bir mimari olarak ta görülür.

Yukarıdaki temellere dayanarak, veri ambarlamayı, verileri yapılandırma ve kullanma işlemi olarak incelersek,

Veri ambarının yapılandırılması, veri tülemeyi, veri temizlemeyi ve verinin birleştirilmesini gerektirir.

Veri ambarının kullanımı ise, sıklıkla iş zekası teknolojilerinin toplanmasını gerektirir. Bu, “bilgi çalışanlarına” (örn:yöneticiler, analistler ve uzmanlar) veri ambarı üzerinden istenilen bilgiyi hızlı ve uygun biçimde elde etmeyi ve bilgiye dayalı sağlam kararları verme imkanı sunar.

“Organizasyonlar veri ambarlarındaki veriyi nasıl kullanıyorlar?”

Pek çok kurum bu bilgiyi iş ile ilgili karar verme aktivitelerinde aşağıdaki bilgileri içeren bir biçimde kullanır:

- Müşterinin ürün seçme ve satın alma modelinin analizini (alım tercihleri, alım zamanı, bütçe döngüleri ve harcama için isteklilik gibi) içeren müşteri odaklılığını artırmak,
- Pazarlama stratejilerinde hassas ayarlamalar yapabilmek için, bölgesel olarak, aylık, üç aylık, yıllık olarak satış performansını karşılaştırarak ürün portföyünün yönetmek ve ürünleri yeniden konumlandırmak,
- Kar sağlayacak kaynakları keşfetmek ve karlılık analizlerini yapmak,
- Müşteri ilişkilerini yönetmek,
- Çevresel düzenlemeler yapmak ve ürün maliyetlerini yönetmek,
- Bütçe planlama, ölçümleme ve değerlendirmelerini yapmak,
- Strateji yönetmek.

2.3. Veri ambarı Özellikleri

W. H. Inmon veri ambarını şu şekilde tanımlıyor: Veri ambarı özneye dayalı, bütünleşmiş, zaman dilimli ve yönetim karar süreçlerini desteklemek amacıyla toplanmış değişmeyen veri koleksiyonudur.[Oracle 10g: Data Warehousing Fundamentals, 2006, pp.2.3]

William Inmon'un 4 kümesi:

Konuya dayalı: Bir veri ambarı, tüketici, tedarikçi firma, ürün ve satış gibi önemli özneler etrafında kurulur. Konuya yönelik olmasının anlamı, veri ambarının kurumdaki yüksek seviyeli varlıklar üzerine odaklanmış olmasıdır. Veri ambarı bir organizasyonun günlük işler ve hareket işleme faaliyetleri üzerinde yoğunlaşmak yerine karar verecek

kimseler için veriye ait modelleme ve analiz üzerinde yoğunlaşır. Özetle veri ambarı veri modelleme ve veritabanı tasarımıyla ilgilenir.

Tümleşik: Bir veri ambarı genellikle farklı veri kaynaklarından çeşitli farklı türde (heterojen) bilgileri bütünleştirerek oluşturulur. Veri temizleme ve veri tümleme teknikleri veri tutarlılığını garantilemek için uygulanır. Farklı veri kaynaklarında verinin tanımlanması açısından değişik yollar kullanılmış olabilir.

Örneğin, müşteri tipi alanı bir uygulamada “B” ve “K” kodlarıyla belirtilmiş olabilir. “B” kodu bireysel, “K” kodu kurumsal müşteri tiplerini belirtmektedir. Bir başka uygulamada müşteri tipi alanı 1 veya 0 değerleriyle ifade edilmiş olabilir. Bir başka uygulama da ise “Bireysel” veya “Kurumsal” ifadeleri müşteri tipi alanını belirtmek için kullanılmış olabilir. Eğer veri ambarında “B” ve “K” kodlarının kullanılması söz konusu ise verinin veri ambarına taşınması esnasında farklı tanımların tümünü bu tanıma göre dönüştürmek gerekecektir.

Zaman Dilimli: Veri ambarları tarihsel veri tutar. Veriler tarihi bir bakış açısından bilgi sağlamak için depolanır. Veri ambarı içerisinde her anahtar yapı zamanın bir elemanı olarak ya kesinlik ya da açıklık içerir. Veri ambarı içerisindeki tüm veri zamanın belirli bir dilimine aittir. Veri ambarındaki veri operasyonel sistemlerde olduğu gibi sadece o andaki değerlerle değil, geçmişteki değerlerle de ilgilenir. Veri zaman içindeki bir noktayla birleştirilerek değerlendirilir. Tarihsel analizler kolaylıkla işlenebilir.

Değişmez: Veri ambarında yer alan verinin değiştirilememesidir. Sadece okunabilir yapıda olmasıdır. Veri ambarları kullanıcıya analiz olanağı verir. Değişen veri yapılan analizin tutarlılığını bozar. Bu yüzden yeni data her zaman mevcuda eklenir güncellenmez.

2.4. Veri Ambarı ve İş Zekası

Veri Ambarı, bir kurumun tüm seviyelerinde karar yapıcı süreçleri destekleyen her çeşit verinin toplandığı stratejik bir veri deposudur. Amaç analitik raporlama, karar verme süreçlerine alt yapı sağlamaktır.

Kurumlar iş zekasını, iş süreçlerini geliştirmek, gözlemlemek, maliyet, kalite ve kontrol amaçlı isterler. Bu bağlamda veri ambarları iş zekasına alt yapı sağlarlar. Bu yüzden iş zekası kurulum projelerinin en önemli ve zor aşamasıdır. Veri ambarının doğru kurgulanması iş zekasının başarı göstergesidir.

2.5. Ambar Verisinin Değişimi - Changing Warehouse Data

Başlangıç veri seti operasyonel sistemden veri ambarına yüklenir. Buna ilk yükleme denir. Bu iş ölçümlerini ve iş analiz kriterlerini içeren veridir. Çoklu kaynak sistemlerden gelen veri yenileme döngüsüne göre veri ambarına eklenir.

Analiz amaçlı kullanılan data değişebilir. Veri ambarında güncel veri tutulmalıdır.

İş birimleri analiz için ne kadar tarihsel veriye ihtiyaç duyduklarını belirler. Örneğin 5 yıl gibi. Daha eski veri arşivlenir veya silinir. Uygunsuz ve/veya hatalı değerler içeren veriler silinmeli veya veri ambarının dışına taşınmalıdır.

2.6. Veri Ambarının Amaçları

Temel olarak veri ambarının amaçları aşağıda ifade edilmiştir:

- Bir organizasyon bilgisini erişilebilir yapar. Veri ambarının içerdiği veriler doğru şekilde etiketlenmiş ve açıktadır. Veriye ulaşmak son derece kolaydır. Organizasyon bilgisini tutarlı yaparlar. Değişik veri kaynaklarından elde edildikleri için tutarlı bilgi veri ambarı için çok önemlidir.
- Uyarlanabilir ve esnek veri kaynağı olmalıdır. Var olan verilerde bir değişiklik yapmadan, yeni veri ve yeni sorguların eklenebilmesine uygun olmalıdır. Teknolojilerin sürekli değişime uygun olarak dizayn edilmesi gerekir.
- Organizasyon verilerin korunması için güvenli olmalıdır. Bir veri ambarı sadece verilerin efektif olarak kontrolünü sağlamaz, ayrıca organizasyona büyük bir bakış açısı, görüş sağlar ve elde edilirse kötü amaçlı kullanılabilir.
- Karar vermeye temel oluşturur. Veri ambarı yönetsel kararlar alabilmek için doğru ve tutarlı verileri sağlar. Yapılan analizler sonucunda alınan kararlar veri ambarlarının bir çıktısıdır. [Kimball et.al. ,1998, pp.10-11].

2.7. Veri Ambarı Bileşenleri

Aşağıda veri ambarı bileşenleri ve içerikleri anlatılmıştır.

2.7.1. Kaynak Sistem (Source System)

Kaynak sistemler aşağıda belirtilen kaynakları içerir :

- Operasyonel Sistemler,
- Arşivler,
- Doğrudan kurumun operasyonel sistemleri ile ilişkilendirilmeyen tablolar (Bireysel hesap tabloları gibi),
- Kurum dışından gelen harici kaynaklar.

Kaynak sistemler ETL süreçleri ile süzülen verilerle Veri Ambarlarını besler.

2.7.2. Data Staging Area

Data Staging Area , verinin veri ambarına taşınmasından önce temizlendiği ve hazırlandığı ortamdır. Bu ortam hem bir veri depo alanı hem de ETL süreçlerinin kurgulandığı bir alandır. İş kullanıcılarına kısıtlanan, sorgulama ve raporlama için uygun bir ortam değildir. Data Staging Area bir veri ambarı uygulamasının gerekliliğidir.

2.7.3. Operational Data Store

ODS, sık sık güncellenen operasyonel datanın bütünleşik bir kopyasıdır. Özellikle OLTP sistemler ve harici sistemlerden yeterli operasyonel raporlama sağlanamazsa, operasyonel raporlamayı güçlendirmek için sıklıkla uygulanır. ODS, veri ambarına operasyonel veri ile beslemek için bir platform sağlar.

2.7.4. Presentation Server

Presentation Area, sorgulama, raporlama ve analiz için verinin depolandığı ve şekillendirildiği bir alandır. Veri ambarı veya bütünleşik veri pazarları dizilerinden oluşur.

2.7.5. Dimensional Model

Boyutlu modelleme, veri modellerinin, işlerin ortak yönleriyle tanımlanan ölçekler kümesi olarak kavramlaştırılması ve görselleştirilmesi için kullanılan bir tekniktir. Özellikle verinin özetlenmesi, yeniden düzenlenmesi ve çözümlenmesini desteklemek üzere, verinin görünümünün sunulması açısından yararlıdır. Boyutlu modelleme, değerler, sayılar, ağırlıklar, bakiyeler ve olaylar gibi sayısal veriler üzerinde yoğunlaşır.

Bu model fact (olay) tablolarını ve dimension (boyutsal) tabloları içerir. Bir fact tablosu tercihen nümerik ve toplanabilir iş ölçümlerinin bulunduğu bir tablodur. Burada 2 veya daha fazla foreign key bulunur, bunlar yardımıyla dimension tablolar fact tabloları ile ilişkilendirilir. [Oracle 10g : Data Warehousing Fundamentals, 2006, pp. 2.19-20]

2.7.6. Data Mart

Datamartlar küçük boyutlu bölümsel ambarlardır. Datamartlar veri ambarının alt kümesidir. Organizasyonun belirli kullanıcıları için ayrılmış ve onlara ait verileri içerir. Bir araya geldiklerinde entegre veri ambarı düzenlenmiş olur. Datamartlar dimension ve fact tablolardan oluşturulmalıdır. Bu şekilde kombine olabilir ve birlike kullanılabilirler.

2.7.7. OLAP (On-Line Analytic Processing)

OLAP, bir kurumun elinde bulunan verileri sadece tek bir bakış açısına göre değil, çok farklı açılardan değerlendirmesine imkan veren bir veri analiz tekniğidir.

OLAP son kullanıcı için veri ambarından sorgulama ve rapor almayı olanaklı kılar. OLAP teknolojisi çok boyutlu Küp (Cube) mantığına dayanır ve OLAP veritabanı çok boyutlu yapıdadır.

2.7.8. Metadata

Veri ambarında verilerin tanımlandığı kısımdır. Metadata depolanmış verinin bilgi ve tanımını tutar. Metadata “Veri hakkında veri” anlamındadır. Metadata her veri elementinin anlamını, hangi elementlerin hangileriyle nasıl ilişkili olduğunu ve kaynak verisi ile erişilecek veri gibi bilgileri içermektedir.

2.8. Veri Ambarı Çatısı (Data Warehouse Framework)

Veri ambarlarının oldukça karışık ve ayrıntılı bir yapısı vardır ve farklı sistemler veri ambarı olarak tanımlanabilmektedir. Bu nedenle tek ve standart bir veri ambarı mimarisinden söz etmek doğru olmayabilir [Ballard et. al., 1998]. Bunun yerine bir veri ambarında bulunması gereken özelliklere değinilmesi ve veri ambarı çatısının incelenmesi daha yerindedir. Bu bölümde, veri ambarlarının bileşenlerinden bahsedilecektir.

2.8.1. Veri Ambarı Mimarisi

Kullanıcıya sunulacak bilginin hangi süreçlerden geçerek hazırlandığı, yani veriler ile verileri üreten ya da kullanan sistemler arasındaki iletişim ve gerçekleştirilen işlemlerin tamamının yapısı anlatılmaktadır.

Bu yapı, birbiriyle ilişkili birkaç bölümden oluşmaktadır [Ballard et. al., 1998]:

- Operasyonel ve Dış Veri Kaynakları Katmanı (Operational And External Systems)
- Bilgi Erisim Katmanı (Information Access Layer)
- Veri Erişim Katmanı (Data Access Layer)
- Metadata Katmanı (Metadata Layer)
- İşlem Yönetim Katmanı (Process Management Layer)
- Uygulama İleti Gönderme Katmanı (Application Messaging Layer)
- Veri Ambarı Katmanı (Data Warehouse Layer)
- Veri Sunum Katmanı (Data Staging Layer)

2.8.1.1. Operasyonel ve Dış Veri Kaynakları Katmanı

Operasyonel sistemler, önemli operasyonel gereksinimleri karşılamak üzere verileri işlerler. Karar faaliyetlerinde doğru ve zamanında bilgi çok önemlidir. Bu bilgi operasyonel işlemler sırasında toplanan veriler içinde mevcuttur. Veri ambarının amacı, kurum içi ve dışı toplanan verileri bilgiye dönüştürmek, operasyonel verilerin içine sıkışıp kalmış olan değerli bilgileri ortaya çıkarmak, başka sistemlerden gelen verilerle birlikte yorumlayarak değerlendirmektir.

2.8.1.2. Bilgi Erişim Katmanı

Veri ambarı mimarisinde Bilgi Erişim Katmanı son kullanıcının katmanıdır. Son kullanıcı araçları olan Excel, Access, SAS gibi araçlar bu katmanda yer alır. Aynı zamanda, statik ve dinamik rapor, çizelge, grafik ya da tabloların üretilmesi, görüntülenmesi ya da çıktılarının alınması gibi işlemler için kullanılan iş zekası araçları da bu katmanda bulunmaktadır.

2.8.1.3. Veri Eriřim Katmanı

Veri Eriřim Katmanı, Bilgi Eriřim Katmanı ile Operasyonel Katman'ın birbiri ile konuřmasını saęlayan katmandır. Veri ambarının en önemli amalarından bir tanesi, uç kullanıcıya “evrensel veri eriřimi” olanaęı saęlamaktır. Evrensel veri eriřimi, uç kullanıcıların Bilgi eriřim aralarından baęımsız olarak kurum ierisindeki verilerin herhangi bir kesimine ya da tamamına eriřebilmesi anlamına gelmektedir.

Bu durumda Veri Eriřim Katmanı, Bilgi Eriřim ve Operasyonel Veri Tabanı katmanları arasında bir arayüz oluřturmalıdır. Bazı durumlarda bu uç kullanıcının gereksinimlerini tam olarak karřılamasına raęmen, kurumlar veri ambarını desteklemek üzere genellikle daha karmařık tasarımlar yaparlar.

2.8.1.4. Veri Kılavuzu (Metadata) Katmanı

Veri Ambarının en önemli katmanlarından biridir. Veri Ambarında verilerin tanımlandıęı katmandır. Metadata kurum iinde veriyi tanımlayan veridir. Metadata her veri elementinin anlamını, hangi elementlerin hangileriyle nasıl iliřkili olduęunu ve kaynak verisi ile eriřilecek veri gibi bilgileri iermektedir.

Veri Ambarındaki veriler, veri ambarı yöneticisinin kullandıęı teknik veriler ve Veri Ambarı kullanıcılarının kullandıęı iř verileri olarak ikiye ayrılır.

Teknik veriler : Operasyonel veritabanı tanımlarını ve veri ambarı tanımlarını ierir. Bu iki tanım veya řema veri ambarını alıřtırılabilmesini saęlayan veri tařıma operasyonlarını ierir. Bu bilgiler Veri Ambarı yöneticisine Veri Ambarında birbiriyle iliřkili verileri gstererek yardımcı olan bilgilerdir.

İř verileri : Kullanıcılara yardım eder. Kullanıcıların veritabanı oluřturan veriler dıřındaki Veri Ambarında bulunan bilgilere ulařmalarına yardımcı olur. Ayrıca veri ambarına verinin ne zaman ve nereden geldięi gibi bilgilerede ulařılmasını saęlar.

2.8.1.5. İřlem Ynetim Katmanı

İřlem Ynetim Katmanı, veri ambarı ve metadata bilgilerinin oluřturulması ve ynetilmesi iin gereken grevleri planlayan katmandır. Veri ambarını gncel tutmak iin gerekleřtirilecek iřlemleri planlayan ve denetleyen iřlemler katmanı olarak dřnlebilir.

2.8.1.6. Uygulama İleti Gnderme Katmanı

Kurum iindeki bilginin tařınmasını saęlayan katmandır. Veri ambarı tařıma sistemidir. Her trl bilgi ve iletileri biriktirerek belirli bir zamanda belirli bir yere gndermek bu katmanın grevlerindedir.

2.8.1.7. Veri Ambarı Katmanı

Veri ambarı katmanı, verinin ilk olarak bilişimsel amaçla kullanıldığı kesimdir. Veri ambarı, verinin mantıksal ya da sanal bir görünümü olarak düşünülebilir; çoğu zaman veri ambarında veri saklanmaz.

2.8.1.8. Veri Sunum Katmanı

Operasyonel ve dış veri kaynaklarındaki bilgilere erişimin sağlanması, veri ambarındaki verilerin seçilmesi, düzenlenmesi, özetlenmesi, birleştirilmesi ve yüklenmesi için gerekli olan işlerin tamamından sorumludur.

Veri kalitesini çözümleyici ve operasyonel veri içerisindeki örüntüleri ve yapıları bulan filtreler de bu katmanda yer alır.

2.8.2. Mimari Seçenekler

Mimarinin seçimi, veri ambarının ve/veya data martın ve yönetiminin nerede olacağını belirleyen ya da bu kararlar doğrultusunda belirlenen bir noktadır. Verinin merkezi bir yerde tutularak merkezi olarak yönetilmesi veya yerel ve/veya uzak yerleşimlerdeki verilerin merkezi bir yerden ya da birbirlerinden bağımsız olarak yönetilmesi gibi. [Gupta, 1997]

Bu bölümde Kurumsal Genişlikte Veri Ambarı (Enterprise), Veri Pazarı (Datamart) ve verinin kaynağına göre sınıflandırabileceğimiz bağımsız (independent), bağlantılı (interconnected) ya da bunların birleşiminden oluşan veri pazarı çeşitleri üzerinde durulacaktır. Sonraki bölümde ise gerçekleştirme seçenekleri olarak yukarıdan aşağıya (top down), aşağıdan yukarıya (bottom up) ya da bu ikisinin birleşimi üzerinde durulacaktır. Ancak unutulmamalıdır ki mimari seçenekler ile gerçekleştirme seçenekleri, bir arada farklı birleşimlerle kullanılabilir. Örneğin, fiziksel olarak dağıtılmış, merkezi olarak yönetilen bir mimaride olup, gerçekleştirilmede ise özel bir çalışma grubuna, bölüme ya da iş koluna hizmet sunan data martlardan başlayan aşağıdan yukarıya gerçekleştirilen bir veri ambarı olabilir.

2.8.2.1. Kurumsal Genişlikte Veri Ambarı (Enterprisewide Data Warehouse)

Kurumun bütün gereksinimleri bir bütün olarak göz önüne alınarak kurumsal veri ambarları tasarlanır ve oluşturulur. Kurumsal veri ambarları kuruluşun genelinde ya da geniş bir kesiminde bulunan analiz verisi için ortak bir havuz gibi de düşünülebilir.

İş analizi için kurumun tüm konu alanlarına ait verileri depolar. Kurumsal veri ambarı, iş konuları hakkındaki tüm bilgiyi, organizasyonun tamamını tarayarak toplar. Kurumsal genişlikte (Enterprisewide) veri ambarları, genelde bir veya daha fazla operasyonel sistemden ve dış bilgi kaynaklarından, veri tülemeyi sağlar. Tipik olarak özetlenmiş veriyi içerdiği gibi detaylı veriyi de içerir ve gigabyte lar seviyesinden, terabyte a veya daha da ötesine uzanan bir dağılıma sahip olabilir. Kapsamlı bir iş modellemesi gerektirir, pahalıdır, tasarımı ile

yapılandırılması çok uzun süre alır ve tüm organizasyon için bilinen bir veri modelinde anlaşma ve elde etmeye ilişkin tutarlılıkta zorluğa bağlı olarak esneklikten yoksundur.

Kurumsal Veri Ambarı kurum içinde fiziksel olarak merkezi ya da dağıtımli olabilir. Dağıtımli Genel Veri Ambarı bütün kurum tarafından kullanılabilmesine rağmen, veriyi kurum içinde fiziksel olarak ayrı yerlerde tutmaktadır.

Dağıtımli yerleşimler başka bir bölüm ya da iş birimi tarafından denetlenebilir. Yani veri ambarına hangi verilerin gönderileceği, veri ambarı güncellenirken başka hangi birim ya da bölümlerin veri ambarına ulaşabileceği, bu birim ya da bölümlerdeki hangi bireylerin erişim hakları olacağı gibi denetimler farklı bölümlerce gerçekleştirilebilir [Boar, 2000].

Veri ambarını oluşturan veri genellikle operasyonel sistemlerden ve dış kaynaklardan, işletim yükünün az olduğu zamanlarda gerçekleştirilen süreçlerle elde edilir. Veri, istenmeyen verilerin elenmesi için filtrelenir, temizlenir, tutarlı hale getirilir, kalite ve kullanılabilirlik gereksinimlerini karşılamak üzere biçimlendirilir ve uç kullanıcıların erişimine sunulmak üzere veri ambarı üzerinde depolanır.

Bir Kurumsal genişlikte Veri Ambarının, kurumsal verinin genel görünümünü sunmaktan daha fazlasını gerçekleştirmesine ve bunun da karar vericiler için çok önemli bir gereksinim olmasına rağmen, bu gerçekleştirmenin hem çok zaman alıcı hem de çok pahalı olduğu bilinmelidir.

2.8.2.2. Veri Pazarı (Datamart)

Data mart, veri ambarının seçilen konular üzerine odaklanan bir bölüm alt kümesidir, bu nedenle alanı bölümsel genişliktedir. Özel seçilmiş olan konularla sınırlandırılmıştır. Veri ambarıyla aynı veri yapısında tanımlanan bir modeldir. Tüm organizasyon veya organizasyonun bir parçasının gereksinimleri için hazırlanırlar. Kullanıcılara bazı avantajlar sağlayan daha az veri içerirler. Hızlı çalışmayı sağlar. Daha az yer işgal eder.

Verinin kaynağına bağımlı olarak bağımsız ve bağımlı olarak sınıflandırılırlar.

2.8.2.2.1. Bağımsız Veri Pazarı Mimarisi (Independent Data Mart Architecture)

Belirli bir İş konusu veya fonksiyonunun ihtiyacı olan verilerin operasyonel ortamlardan alınarak daha kolay raporlanabilir ve analiz edilebilir hale getirilmesi sonucunda oluşan veritabanlarıdır. [Oracle 10g: Data Warehousing Fundamentals, 2006, pp.2.18]

Bağımsız veri pazarı mimarisi, bir veya daha fazla operasyonel sistemden, dış bilgi kaynaklarından, belirli bir konu veya kısmi bir bölüm içerisinde yerel olarak oluşturulmuş veriden kaynak olarak yararlanır.

Bağımsız Veri Pazarı Mimarisi, özel bir çalışma grubu, bölüm ya da iş birimi tarafından yönetilen ve sadece bu birimlerin kendi gereksinimlerini karşılamak üzere

oluşturulan bağımsız veri pazarlarını belirtmektedir. Bu veri pazarlarının diğer birimlerdeki veri pazarlarıyla bir ilgisi bulunmayabilir. Her veri pazarının verisi birim içerisinde üretiliyor olabilir. Verinin eğer varsa bir kurumsal veri ambarından da gelmesi olanaklıdır.

Yapının tümleşik olmaması ve verinin genel görünümünün elde edilememesi gibi kısıtlar söz konusudur. Çünkü herhangi bir birime ait veri pazarındaki veriye, yalnızca veri pazarının sahibi olan birim tarafından erişilebilmektedir.

2.8.2.2.2. Bağımlı Veri Pazarı Mimarisi (Interconnected Data Mart Architecture)

Veri Ambarından belirli bir İş grubunun veya fonksiyonunun ihtiyacı olan verilerin alınması ve daha kolay raporlanabilir ve analiz edilebilir hale getirilmesi sonucunda oluşan veritabanlarıdır. [Oracle 10g: Data Warehousing Fundamentals, 2006, pp.2.17]

Bağımlı Veri Pazarı Mimarisi temel olarak dağıtım bir yaklaşımdır. Veri pazarları verinin kurum genelindeki görünümünü elde etmek amacıyla bütünleştirilebilir ya da birbirlerine bağlanabilir. Bütünleştirme sonucunda veri pazarları Kurumsal Veri Ambarı haline gelebilir. Böylece kurumdaki kullanıcılar, diğer bölümlere ait veri pazarına erişebilir ve bunu kullanabilir. Bu mimaride, veri operasyonel ya da dış veri kaynaklarından geliyor görünmesine rağmen, eğer varsa bir kurumsal veri ambarından da gelmesi olanaklıdır.

Bu mimari, birçok işlev ve yeteneği de beraberinde getirmektedir. Ancak unutulmamalıdır ki bu yeni seçeneklerle birlikte bütünleştirme gereksinimleri ve karmaşıklık da artacaktır. Örneğin bu durumda ortamın kim tarafından yönetildiği ve denetlendiğinin bilinmesi, birden fazla veri pazarı tarafından ortak kullanılan verileri tutmak üzere yeni bir katman gereksinimi ya da veri pazarları arasında verinin paylaşılması için kullanılacak bir plan hazırlanması gibi durumlar düşünülebilir. Bunların her biri, mimarinin karmaşıklığını arttıracaktır.

Bağımlı Veri Pazarları, birimler tarafından birbirlerinden bağımsız olarak denetlenebilir. Veri pazarına hangi verinin yükleneceğine, verinin ne zaman günleneceğine, veriye kimin erişeceğine ve verinin nerede tutulacağına bu birimler karar verir. Böylece Bilişim Sistemleri bölümünün kaynaklarına olan gereksinim de çok azalacaktır. Bilişim Sistemleri bölümü ise gerçekleştirmenin kurum içi güvenlik, yedek alma/kurtarma ve ağ bağlantıları gibi konularda yardım sunabilir. Aksi düşünülürse, veri pazarlarının yönetilmesi bu bölüm tarafından da üstlenilebilir. Her birimin kendisine ait veri pazarı olmasına rağmen, veri pazarlarının gerçekleştirilmesi için gerekli araçlar, çalışanlar ve kaynaklar Bilişim Sistemleri Bölümü tarafından sağlanabilir.

2.8.3. Veri Ambarı Geliştirme Yaklaşımları

Bunlardan bazıları yukarıdan aşağıya, aşağıdan yukarıya ve ikisinin birlikte olduğu seçenekleridir.

Gerçekleştirme yöntemi seçimi, bilgi sistemi altyapısı, uygun kaynak miktarı, seçilen mimari, uygulamanın kapsamı, kuruluştaki genel veriye duyulan gereksinim, yatırım getirileri ve uygulamanın hızı gibi birçok etkenden etkilenebilir.

2.8.3.1. Yukarıdan Aşağıya Yaklaşım

Yukarıdan aşağıya 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. Bu da veri ambarı uygulamasında yer alacak her çalışma grubu, bölüm ve görevden bir kişinin projeye katılması gereksinimini doğurur. Gerçekleştirim başlamadan önce, kullanılacak veri kaynağı, güvenlik, veri kalitesi, veri standardı ve genel veri modeli ile ilgili kararların verilmiş olması gereklidir.

Veri pazarları, doğrudan işlevsel ya da dış veriler yerine kurumsal veri ambarı kullanılarak çoğaltılırlar.

Yukarıdan aşağıya gerçekleştirim daha tutarlı veri tanımlamaları oluşmasına neden olur ve iş kurallarının kurum genelinde geçerli olmasını sağlar. Ancak başlangıçtaki planlama ve tasarım maliyetleri yüksek olabilir. Planlama ve tasarım, zaman alıcı bir süreçtir. Bu nedenle gerçekleştirmeyi ve yatırım getirilerini geciktirebilir. Farklı çalışma grupları, bölümler ve görev tanımları arasında ortak veri tanımlamaları ve iş kuralları oluşturmak zaman alan ve zor bir iştir. Kurumsal bir veri modeli geliştirmek de uzun zaman isteyen bir süreçtir. [Oracle 10g: Data Warehousing Fundamentals, 2006, pp.2.24]

2.8.3.2. Aşağıdan Yukarıya Gerçekleştirme

Aşağıdan yukarıya gerçekleştirim, deneyler ve prototipler ile başlar ve veri pazarlarının daha genel bir yapının kullanıma geçmesini beklemeden planlanmasını ve tasarımını içerir. Bu durum iş modellemenin ve teknoloji gelişiminin erken aşamalarında kullanışlıdır. Bir organizasyona önemli taahhütlerde bulunması öncesinde teknolojinin yararlarını değerlendirmeyi sağlar ve çok az bir masrafla organizasyonun ileri gitmesine imkan verir.

Ancak bu, genel bir yapının geliştirilmeyeceği anlamına gelmez. Veri pazarları geliştikçe, genel yapı da şekillenir. Bu yaklaşım yukarıdan aşağıya gerçekleştirim yaklaşımından daha çok kabul görmüştür. Bunun nedeni, veri pazarlarından anında elde edilen verinin genel yapıyı şekillendirmede ve düzenlemede kullanılabilir olmasıdır. Yukarıdan aşağıya gerçekleştirim aksine, veri pazarları genel veri ambarına paralel olarak

geliştirilebilir. Veri pazarları genel veri ambarından ya da doğrudan işletimsel ya da dış veri kaynaklarından çoğaltılabilirler.

Birçok kurum, özellikle iş yönetimi alanında, proje hazırlıklarının daha çabuk ve masrafsız yapılabilmesini sağladığı için, aşağıdan yukarıya gerçekleştirme yaklaşımını benimsemiş durumdadır. Veri pazarları, genel bir veri ambarından daha az karmaşık oldukları için daha hızlı sonuçlar üretme olanağı verir. Bununla birlikte, başlangıçtaki basit gerçekleştirme, gerek donanım, gerek diğer kaynaklar açısından, genel bir veri ambarı konuşlandırma sürecinden daha az masraflıdır.

Aşağıdan yukarıya gerçekleştirme yaklaşımının, göz önünde bulundurulması gereken bazı olumsuz yanları da vardır;

Veri pazarlarının sayısı arttıkça, gereksiz veri ve tutarsızlık miktarı artabilir. Dikkatli bir planlama, izleme ve tasarımla bu miktar en aza indirilebilir. Birden çok veri pazarı, daha fazla veri çıkarma işlemi gerektirdiğinden, operasyonel sisteme artan bir yük bindirecektir. Veri pazarlarının daha genel bir yapıda bütünleştirme istendiğinde, bir dereceye kadar planlama yapmak gerekmektedir. Gerçekleştirme genişledikçe, aynı işler tekrarlanabilir. Gerçekleştirmenin mevcut kesimlerinin değiştirilmesini gerektiren yeni sorunlar ortaya çıkabilir. [Oracle 10g: Data Warehousing Fundamentals, 2006, pp.2.25]

2.8.3.3. Birleştirilmiş Yaklaşım

Görüldüğü gibi aşağıdan yukarıya ve yukarıdan aşağıya yaklaşımların artıları ve eksileri bulunmaktadır. Birçok durumda her ikisinin birleştirildiği bir yaklaşım en iyi sonucu verir. Birleştirilmiş gerçekleştirimde organizasyon, tabandan tepeye yaklaşımının hızlı uygulaması ve fırsatlar sunan kullanımına sahipken, yukarıdan aşağıya yaklaşımının planlı ve stratejik özelliğini de kendi yararına kullanabilir. Her ne kadar zor bir dengeleme gerektirse de, bu iş iyi bir proje yöneticisiyle başarılabilir. Bu yaklaşımın temel noktalarından birisi, aşağıdan yukarıya gerçekleştirme sırasında veri pazarları kuruldukça tümleştirilmelerini destekleyecek genel yaklaşım için gerekli olan planlama ve tasarım derecesinin belirlenmesidir.

İlk adım olarak gerçekleştirmeye katılacak olan görev grupları belirlenir. İş sürecinin genel görünüşü ve ilgilenecekleri veri alanları, veri pazarlarının geliştirilmesi planı için önemli veri öğelerini sağlar. Veri pazarları oluştuğunda, birçok veri pazarı tarafından gereksinim duyulan veri öğelerinin ne şekilde ele alınacağına dair bir plan oluşturulur. Bu plan, daha genel bir veri ambarı yapısı için ya da bütün veri pazarlarının ulaşabileceği basit bir veri marketi için başlangıç noktası oluşturabilir. Bazı durumlarda farklı veri ambarları arasındaki verileri çoğaltmak uygun olabilir. Veri depolama alanını miktarı en az düzeyde tutulması ve

veriye kolay erişim sağlanırken, veri pazarları arasında veri tutarsızlıklarının ve gereksiz veri depolamanın da olabildiğince engellenmesi gerekir.

Bu sorunlar, her veri ambarı uygulamasında çözümlenmelidir. Birleştirilmiş yaklaşımı kullanmak karşılaşılan sorunları genel veri ambarı düzeyinde değil, daha küçük kapsamda, yani veri pazarı düzeyinde çözebilme olanağı sağlar. Gerçekleştirme sürecinin dikkatle izlenmesi ve sorunların giderilebilmesi ile her iki yaklaşımın faydalarından en üst düzeyde yararlanılabilir.

2.9. Verinin Biçimlendirilmesi

Veri ambarı, farklı grup kullanıcılar arasında karar verebilmeyi sağlayan, konuya yönelik, tümleşik ve zamanla değişken veri derlemidir. Amaç, veriyi filtreleyip, uygun şekillere dönüştürüp, özetleyip derleyerek kullanıcılar tarafından kolayca erişilip çözümlenecek yapılarda sunmaktır. Bu yapı, veri ambarı gerçekleştirmede öncelikle tanımlanmalıdır. Veri, özelliklerine göre biçimlendirilip yerleştirilir.

Veri ambarı veri modeli geliştirmede, kurumsal bir veri modeline (Enterprise Data Model – EDM) sahip olmak yararlı olabilir, ancak zorunlu değildir. Kurumsal veri modelinden, iş gereksinimleri ile ilgili genel bir bakış açısı ve bilgi edinilebilir. Bununla birlikte kurumsal veri modeli, özel bir ilgi alanı ile ilgili fiziksel tasarım ve veri öğeleri ile bağlantı kurabilmeyi sağlar.

Verinin biçimlendirilmesinde verinin duyarlılığı en önemli ölçütlerden biridir. Yüksek duyarlılıktaki veri her sorguyu destekleyebilir. Ancak işlenmesi ve yönetilmesi gereken yoğun veri, işlem zamanını arttırarak yanıt sürelerini uzatır. Düşük duyarlılıkta veri ise sadece belirli sorguları destekleyebilir. Ancak veri miktarının azalmasıyla başarımlar belirgin bir şekilde artacaktır.

Veri ambarının boyu değişkendir ve genelde çok büyüktür. Geçmişe yönelik verinin de saklandığı düşünüldüğünde, böyle olmasının doğal olduğu görülür.

2.10. Verinin Yapılandırılması

Veri ambarı oluşturulurken, bir kuruluşun gereksinimlerini karşılamak için kullanılan 3 temel veri türünden bahsedilebilir:

- Gerçek zamanlı veri (Real-time data)
- Türetilmiş veri (Derived data)
- Uzlaşılmış veri (Reconciled data)

Bu bölümde, kullanım şekli, kapsamı ve geçerlilik durumuna göre verinin 3 türü tanıtılacaktır.

2.10.1. Gerçek Zamanlı Veri

Gerçek zamanlı veri işin o anki durumunu belirtir. Genellikle operasyonel uygulamalar tarafından işleri yürütmek için kullanılır. Hareket oldukça veri de aynı paralellikte sürekli değişir. Gerçek zamanlı verinin duyarlılığı yüksektir.

Sadece operasyonel sistemlerle sınırlı olmayan gerçek zamanlı veri, kurum içerisinde bilgi sistemlerine de dağıtılabilir. Gerçek zamanlı verinin operasyonel, yönetim ve taktik kararlar için oldukça önemli olduğu bankacılık sektöründe, bağımsız bir sistem, çözümleme yapılması ve stratejik kararların verilmesi için veriyi operasyonel sistemden bilgi sistemlerine (veri ambarı) aktarır. Gerçek zamanlı veriyi veri ambarında kullanabilmek için verinin uygun kalitede (belki özetlenerek) çözümleyiciler tarafından anlaşılabilir ve işlenebilir bir duruma dönüştürülmesi gerekir. Bunun nedeni gerçek zamanlı verinin operasyonel sistemin kullanacağı ve süzülmesi gereken verinin yanısıra her harekete ilişkin detaylı veriyi de içeriyor olmasıdır. Bununla birlikte birçok farklı sistemden geliyor olması nedeniyle tutarlı olmayabilir. Örneğin para birimleri ve kur değerleri farklı sistemlerde farklı olabilir. Bu farklılıklar, gerçek zamanlı veri veri ambarına yüklenmeden önce giderilmelidir.

2.10.2. Türetilmiş Veri

Türetilmiş veri, gerçek zamanlı verinin, bir takım işlemler sonucunda özetlenmiş, ortalaması alınmış veya bir araya getirilmiş halidir. Türetilmiş veri gereksinime göre detaylı ya da özet halde bulunabildiği gibi, işin belirli bir zamandaki ya da geçmişte belirli bir zaman aralığındaki durumunu gösteren kayıtları olabilir.

Türetilmiş veri çözümleme ve karar verme aşamalarında kullanılır. Çözümlemede nadiren büyük oylumlu ayrıntılı veriye gereksinim duyulur. Genellikle işleme ve kullanması kolay özet veri kullanılır. Büyük oylumlu atomik veriyi işlemek, büyük oranda işlem gücü kaynağı gerektirir. Sorgu işleme kapasitesinin iyileştirilmesi gereksinimi göz önüne alındığında, en etkin yaklaşım, türetilmiş veri öğelerinin önceden hesaplamak ve ayrıntılı veriyi gereksinimleri en iyi şekilde karşılayacak şekilde özetlemektir. Büyük oylumlu verinin uygun sürelerde işlenebilmesi, çözülmesi gereken en önemli sorunlardan birisidir.

2.10.3. Uzlaşmış Veri

Uzlaşmış veri, artırılmış, uygun şekilde ayarlanmış ya da veri çözümleyicileri tarafından kullanılacak tümleşik ve kaliteli bir şekilde geliştirilmiş gerçek zamanlı veridir. Kaliteli olmasının temel koşulu tutarlı olmasıdır. Bu işlemler sırasında geçmişe dönük veri de korunabilir. Uzlaşmış veri, aslında türetilmiş verinin özel bir türüdür.

Nadiren açıkça tanımlanan uzlaşmış veri, genellikle t retme iřleminin mantıksal bir sonucu olarak ortaya ıkar. Bazı durumlarda ise operasyonel veriye d n řt r lmek  zere geici k t klerde saklanır.

2.11. Veri Ambarı iin Veri Modelleme

Veri ambarı, veri  z mlleme ve karar alma s relerinde kullanılacak t mleřik ve tutarlı verinin saėlanabilmesi iin en uygun yaklařım olarak kabul edilmektedir. Ancak veri ambarlarının gerekleřtirilmesi karmařık, ok zaman ve kaynak gerektiren bir iřtir. Bu durum  zellikle b y k kurumlar iin geerlidir. Kazanımların hızlıca elde edilebilmesi amacıyla, “veri pazarlarının yukarıdan-ařaėıya gerekleřtirilmesi” seeneėi  n plana ıkmıřtır. K  k aptaki bu t r gerekleřtirmeler, yatırımın kısa vadede geri d n ř m n  saėlamaktadır. Veri pazarları, genel veri ambarlarının gerekleřtirilmesine de engel olmamakta, b y t lmek ya da birleřtirilmek yoluyla kurumun genel veri ambarı  z m ne katkıda bulunmaktadır. Veri ambarlarının genel veri ambarı ya da veri pazarı olarak gerekleřtirilmesi ise  nemli getirileri olduėu gereėini deėiřtirmemektedir.

Burada sorulması gereken soru, veri ambarı kullanıcılarının gereksinimlerini en iyi Őekilde karřılayabilmek amacıyla veri tabanlarının nasıl geliřtirileceėi; bir bařka deyiřle, verinin nasıl modelleneceėidir. Veri modeli, veri  zerinde gerekleřtirilecek b t n iřlemlerin bir parasıdır.

Bu b l mde bir veri ambarının gerekleřtirilmesi amacıyla verinin nasıl modelleneceėi  zerinde durulacak, iki temel veri modelleme tekniėinden s z edilecektir: Varlık Baėıntı Modelleme (Entity Relationship (ER) Modeling) ve Boyutlu Modelleme (Dimensional Modeling).

Operasyonel ortamda aėırlıklı olarak kullanılmakta olan VB modelleme tekniėi, veri ambarcılıėıyla birlikte veri  z mlleme ortamını destekleyen bir teknik olarak ortaya ıkmıřtır. Ancak VB modellemenin saėladıėı  nemli desteėe raėmen, veri ambarcılıėında kullanılacak modelleme tekniėi olarak boyutlu modelleme tekniėi ilgi odaėı haline gelmiřtir.

2.11.1. Veri Modelleme Teknikleri

Bir veri ambarı ortamında kullanılan iki farklı veri modelleme tekniėi Varlık Baėıntı Modelleme ve Boyutlu Modelleme teknikleridir.

2.11.1.1. Varlık-Baėıntı Modelleme

VB modelleme kavramı, veri tabanı tasarımıyla uėrařmıř kimseler iin bildik bir kavramdır. VB modelleme, ilgilenilen alanın veri modelini iki temel kavrama dayanarak oluřturur:

Varlıklar (entity) ve varlıklar arasındaki baėıntılar (relationship).

Ayrıntılı VB modelleri, varlıklara ya da bağıntılara ait olan nitelikleri de içerebilir. İş dünyası ve karmaşık sistem ortamları arasındaki net olmayan bağıntıların basitleştirilmesini ve anlaşılmasını sağladığı için, VB Modeli bir soyutlama aracı olarak görülebilir.

Varlık-bağıntı veri modeli , veritabanı şeması bir varlık kümesi ve onların arasındaki ilişkilerden oluştuğundan, ilişkisel veritabanı tasarımında yaygın olarak kullanılmaktadır

2.11.1.2. Boyutlu Modelleme

Veri ambarları ve OLAP araçları çok boyutlu veri modeli üzerine temellenmiştir. Bu model, veriyi veri küpü formunda incelemektedir. Boyutlu modelleme, kullanıcının veri tabanı tabloları bağlamındaki gereksinimlerini güçlü bir şekilde ifade edebilmektedir.

Boyutlu modelleme, veri modellerinin, işlerin ortak yönleriyle tanımlanan ölççekler kümesi olarak kavramlaştırılması ve görselleştirilmesi için kullanılan bir tekniktir. Özellikle verinin özetlenmesi, yeniden düzenlenmesi ve çözümlenmesini desteklemek üzere, verinin görünümünün sunulması açısından yararlıdır. Boyutlu modelleme, değerler, sayılar, ağırlıklar, bakiyeler ve olaylar gibi sayısal veriler üzerinde yoğunlaşır.

Boyutlu modellemede temel kavramlar :

- Olaylar,Olgular (facts)
- Boyutlar (dimension)
- Ölçekler (measure) / değişkenler (variables)

2.11.1.2.1. Olay

Olaylar rakamsal ölçülerdir. Boyutlar arası ilişkileri analiz etmek için kullanacağımız miktar bilgileri olarak tanımlanabilir.

Bir olay, ölççeklerden ve bağlam(anahtar) verilerinden oluşan ilişkili veri öğelerinin bir derlemidir. Her olay tipik olarak bir iş ögesi, iş hareketi ya da iş sürecini çözümlenmekte kullanılabilen bir oluşu ifade eder. Bir veri ambarında, olaylar sayısal verinin saklandığı ana tablolarda gerçekleştirilmektedir.

Olay tablosu, olguların isimlerini veya ölçü-miktarlarını ve ilişkili boyut tablolarının her birine ait anahtar içerir.

2.11.1.2.2. Boyut

Boyut, aynı türdeki görüntülere ait üye ya da birimlerin bir derlemidir. Bir boyut bir çizelgede genel olarak bir eksene karşılık gelir. Boyutlu modelde, olay tablosundaki her veri noktası, çok sayıdaki boyutların her birinden bir ve yalnızca bir üyeye ilişkilendirilir. Yani boyutlar olaylar için kavramsal altyapıyı sağlarlar. Birçok çözümlenme işlemi olaylar üzerindeki boyutların etkisini sayıya dökmek üzere kullanılmaktadır.

Boyutlar, üzerinde Çevrimiçi Çözümlenme İşlemi (Online Analytical Processing –

OLAP) yapılacak olan parametrelerdir. Örneğin, bütün ürünlerin satışlarını çözümlmek için kullanılan bir veri tabanındaki boyutlar şunlar olabilir:

- Zaman
- Yer/bölge
- Müşteriler
- Satıcılar
- Gerçek, bütçelendirilmiş ya da tahmin edilen sayılar gibi senaryolar

Boyutlar genellikle branş ya da çalışanlar gibi, sayısal olmayan, bilgi verici varlıklarla eşleştirilebilir.

Boyut Üyeleri : Bir boyutta bir çok boyut üyesi vardır. Boyut üyesi, bir veri ögesinin yerini belirlemek üzere kullanılan ayrı bir isim ya da belirleyicidir. Örneğin, bütün aylar ya da yıllar bir zaman boyutu oluştururken, bütün şehirler, bölgeler ya da ülkeler de bir coğrafya boyutu oluşturmaktadır.

Boyut Sıradüzeni (Hiyerarşi) : Boyut üyeleri bir ya da daha fazla sıradüzende düzenlenebilir. Herbir sıradüzenin de birden fazla düzeyi olabilir. Bir boyutun her üyesi aynı sıradüzen yapısında yer almaz. Şekil 2.1'deki zaman boyutu sıradüzeni güzel bir örnek olarak ele alınabilir. Zaman boyutu için iki sıradüzen tanımlanmasının nedeni, bir haftanın 2 aya, 3 aylık dönemler arasına ya da daha üst düzeylere yayılabilir olmasıdır. Bu nedenle ayları.. vs elde etmek için haftalar toplanamaz. Dolayısıyla haftalar için ayrı bir sıradüzen tanımlanmıştır. Eğer verinin haftalar düzeyinde çözümlenmesinin pratikte bir faydası yoksa, ikinci sıradüzeni oluşturmaya gerek yoktur.

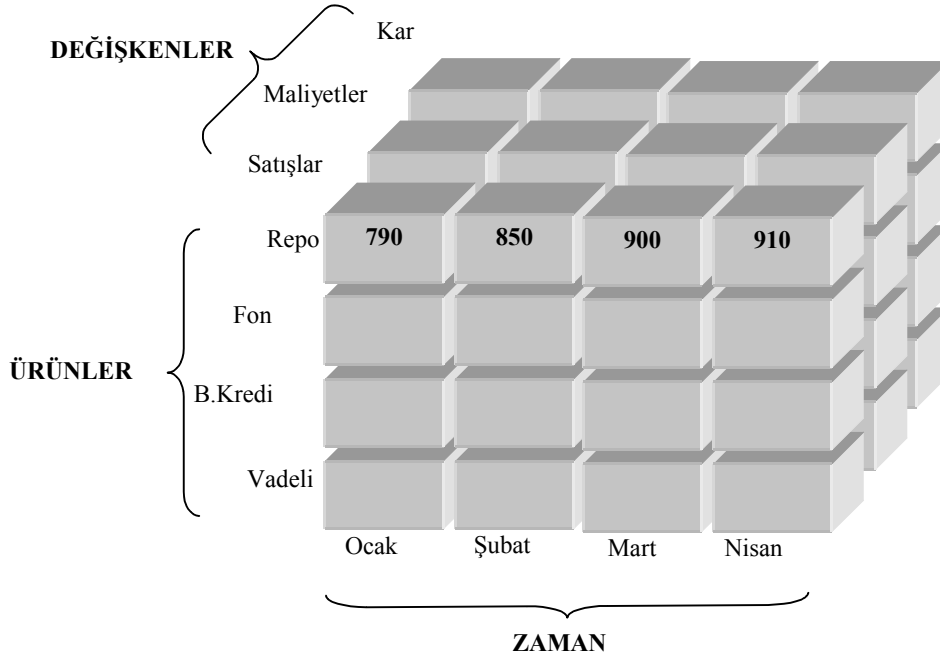
2.11.1.2.3. Ölçek

Ölçek, bir olayın sayısal niteliğidir. Bu nitelik, işin boyutlara göre başarımını ya da gidişatını gösterir. Sayılar değişken olarak adlandırılır. Örneğin satışların parasal miktarı, satış hacmi, hareket miktarı gibi değerler ölçek olarak düşünülebilir. Ölçekler, boyut üyelerinin birleşimleriyle belirlenir ve olaylar üzerine yerleştirilir.

2.11.1.2.4. Küp

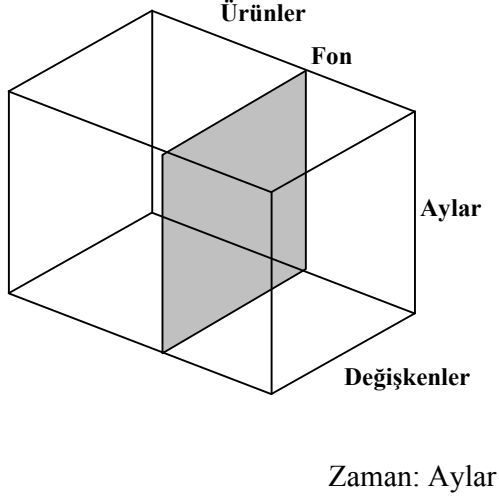
OLAP'ın sunduğu en büyük avantaj, veriler üzerinde sağladığı çok yönlü bakış imkanıdır. Örnek olarak bir ilişkisel veri tabanı iki boyutlu bir ortam olarak düşünülebilir. Bilgi, satır ve sütunlarla temsil edilmektedir. Buna karşın bir uygulamada Satışlar, Maliyetler ve Kar profilleri ele alınarak üç boyutlu bir analiz gerektiğinde OLAP sistemlerinin kullanımı gerekecektir. (Şekil 2.1). OLAP, çeşitli boyutlardan aldığı verileri bir platformda birleştirilmiş gibi geniş bir boyutlama sağlayacaktır.

OLAP'ı sadece büyük özet tablolar gibi yorumlamak doğru olmaz. Excel kullanıcıların tanıdıkları pivot tabloların, çok gelişmiş ve hızlı bir hali olarak düşünülebilir. Tasarlanan bir OLAP yapısının, hiyerarşilerini ve boyutlarını görmek mümkün olsa da, verileri nasıl tuttuğunu, nasıl sorgulanacağını sadece mdx sayesinde görebiliriz. Fakat dışarıdan baktığımızda iç içe geçmiş küpler olarak yorumlayabiliriz. Bu nedenle OLAP yapılarına, "küp" adı verilmiştir.



Şekil 2.1. Üç Boyutlu Küp

OLAP kavramının anlatılmasında, "küp" kavramından yararlanılır. Veriler boyuta göre düzenlenirken, küpün kenarlarına yerleştirilir. Küpün her bir kenarı, bir boyuta karşılık gelir (örneğin, satış boyutu, ürün boyutu, müşteri boyutu, zaman aralığı boyutu, vb.). Boyutlanmış olan veri, zar şeklinde parçalara bölünerek ve birbirine eklenerek, istenilen boyut birleşimlerinde gözlemlenebilir.



Ürün: Fon

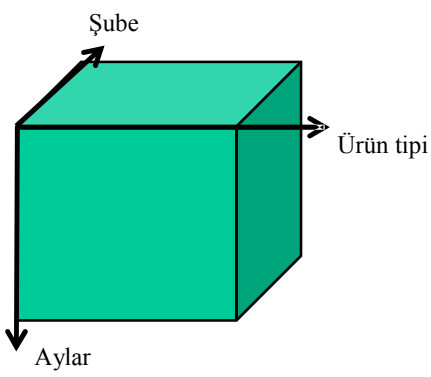
Değişkenler: Tümü

	Satışlar	Maliyetler	Kar
Ocak	600	500	100
Şub.	500	450	50
Mart	450	375	75
Nis.	475	350	125
May.	500	420	80
Haz.	425	250	175
Tem.	380	230	150
Ağus.	480	300	180
Eylül	500	300	200
Ekim	400	220	180
Kasım	500	310	190
Aralık	650	400	250

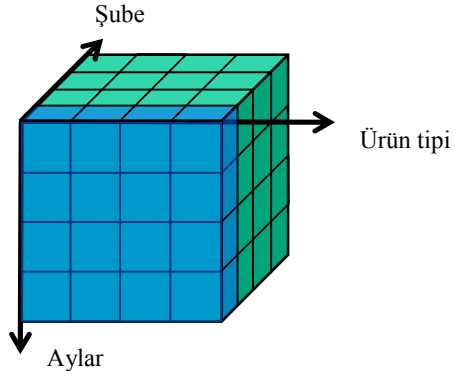
Şekil 2.2. Küpün bir zarı (Fon)

OLAP araçlarının önemli bir özelliği de daha zor sorgulamaların gerektirdiği karmaşık matematiksel işlemleri ve analizleri oldukça hızlı biçimde gerçekleştirebilmeleridir. İşletmeler, iş performansları ile ilgili ölçütlere ulaşarak rekabet avantajını elde etmeye çalışırlar ve finansal ve operasyonel performansı geliştirmede ihtiyaç duyulan bilgiye sürekli olarak ulaşmak için çalışırlar. OLAP; değişen müşteri istekleri ve operasyonel ihtiyaçlara, gelişen pazara ayak uydurmayı sağlayacak sürekli ve hızlı bilgi kaynağı sağlar.

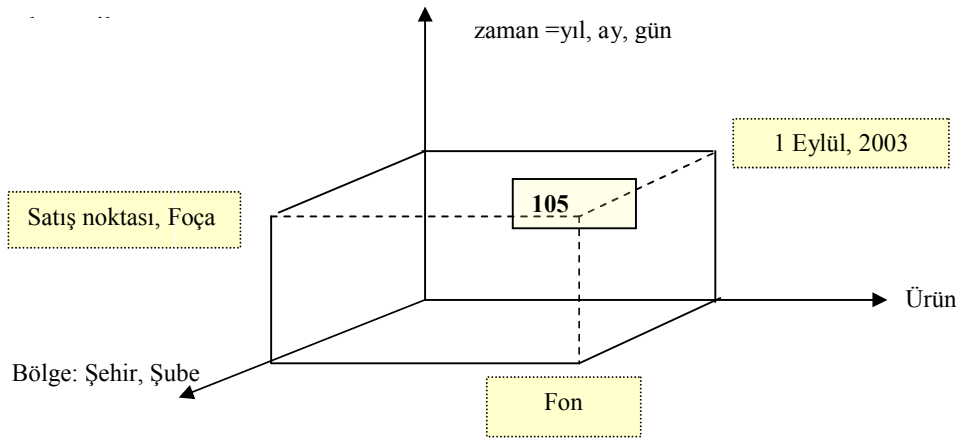
Veriler bir küp üzerinde gösterilir. Üç boyutlu genel bir küp aşağıdaki şekilde tanımlanabilir. Boyutlardan biri ürün tipini, diğer boyut, ürünün satıldığı şubeyi, diğer boyuttada ürünün hangi ay satıldığını temsil eder. Küpün herhangi bir hücrenin üzerindeki sayısal değer de satılan birimler, maliyetler gibi herhangi bir ölçüyü ifade eder. Bu küp üzerinden herhangi bir zar alındığında bu zar küp üzerinde tek bir durumu gösterir. Örneğin Yalnızca Ankara şubesindeki satışlar gibi. Bu zar aslında iki boyutlu bir durumu gösterir.



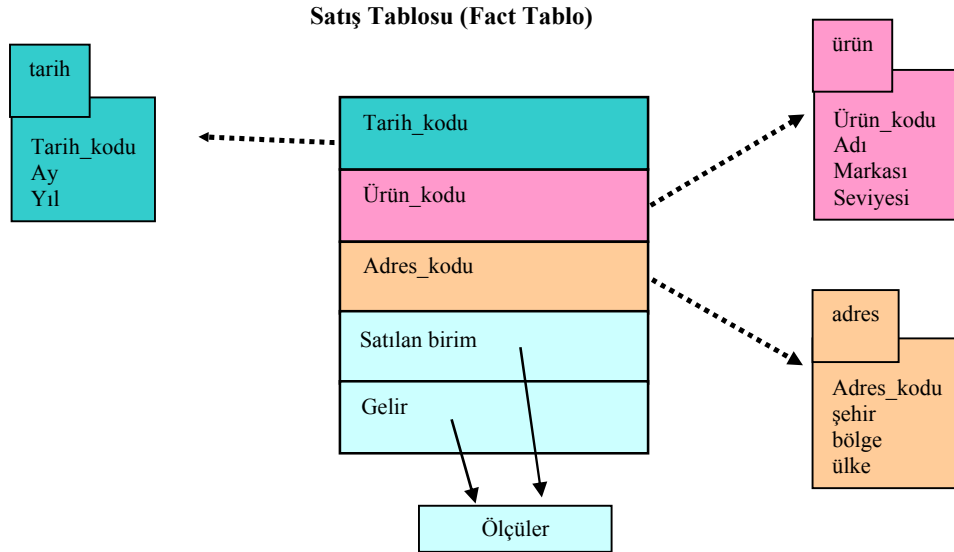
Şekil 2.3. Üç boyutlu bir küp



Şekil 2.4. Üç boyutlu küpten alınan bir zar(3D>2D)



Şekil 2.5. OLAP verisini temsil eden veri küpü



Şekil 2.6. Satış analizi için üç boyutlu bir verinin ilişkisel gösterimi

Foça şubesinin neden Konak şubesinden daha az satış yaptığını araştırmak istediğimizde ise Drill-Down işlemi gerçekleştirmemiz gerekir. (Şekil 2.8)

Satış Hacmi (x Milyon TL)		Fon		Repo	
		A	B	1 Gün	7 gün
İzmir	Foça	33	12	8	12
	Konak	45	34	20	23



Drill-Down
Boyut: Yer
Üye: Foça

Satış Hacmi (x Milyon TL)		Fon		Repo	
		A	B	1 Gün	7 Gün
Foça	1. Portföy	20	8	6	7
	2. Portföy	13	4	2	5

Şekil 2.8. Drill-Down örneği

2.11.1.2.5.2. Slice ve Dice (Dilimleme ve Çevirme)

Slice ve dice, görselleştirilmiş küpteki veriler üzerinde dolaşma işlemleridir. Slice işlemi, kullanıcıların bazı özel bakış açılarından odaklanabilmeleri için küpü dilimlere ayırır. Dice ise kullanıcıların veri çözümleme işlemi daha özel açılardan yapabilmeleri için küpü farklı bakış açılarına göre döndürür. Belirli bir ay için şube ve ürünlere göre oluşturulmuş bir raporu incelemek amacıyla elimizde 3 aylık dönemlerdeki toplam satışın şube bazında miktarları olsun. Boyutun üründen zamana çevrilmesi işlemi dice işlemi olacaktır. Toplam satış yerine sadece fon satışına odaklandığımızda, küpü aynı boyutlardan sadece fon için kesebiliriz. Bu ise slice işlemine karşılık gelmektedir.

2.11.1.2.6. Çok Boyutlu Veritabanları için Şemalar

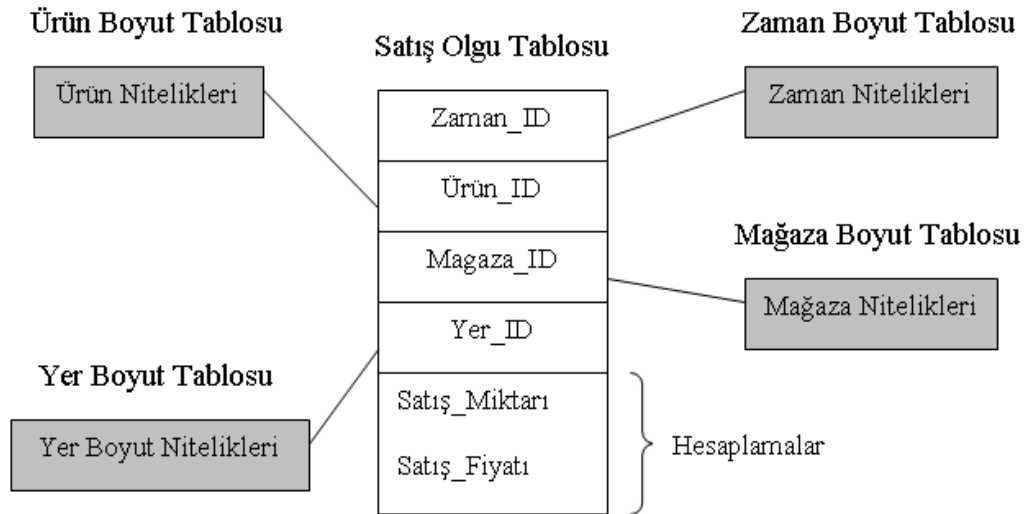
Varlık-bağıntı veri modeli , veritabanı şeması bir varlık kümesi ve onların arasındaki ilişkilerden oluştuğundan, ilişkisel veritabanı tasarımında yaygın olarak kullanılmaktadır. Böyle bir veri modeli çevrim içi hareket işleme için uygundur.

Bir veri ambarı bununla birlikte çevrim içi veri analizini kolaylaştıran kısa, konuya dayalı bir şema gerektirir.

Veri ambarı için en popüler veri modeli, çok boyutlu modeldir. Böyle bir model yıldız şema (Star), kar tanesi şema (Snowflake) ve olgu takımyıldızı (constellation) şeması biçiminde olabilir. Bütün şema tiplerini inceleyelim.

Yıldız şema : En çok bilinen modelleme örneği, içerisinde veri ambarının içerdiği en önemli veri kısmını gereksiz fazlalık olmadan içinde bulunduran büyük bir merkezi tablo (olay/olgu (Fact) tablosu) ve her biri bir boyut için olmak üzere küçük yardımcı tablolar kümesi (boyut tabloları (dimension)) bulunduran yıldız şemadır. Şema çizgesi, merkezi olay tablosunun etrafında merkezden çıkan bir desen içerisinde gösterilen boyut tabloları ile, starburst yapısına benzer. [Oracle 10g: Data Warehousing Fundamentals, 2006, pp.3.6]

Yıldız modeli, boyutlu model için temel yapıdır. Genellikle bir büyük merkezi tablo (olgu tablosu denir) ve olgu tablosu etrafında yıldız biçiminde sıralanmış daha küçük tablolar söz konusudur. Şekil 2.9'da, merkezde bir olgu tablosu ve olgu tablosu (Satış) etrafında sıralanmış zaman, mağaza, yer ve ürün boyut tablolarını gösteren bir yıldız modeli örneği görülmektedir.



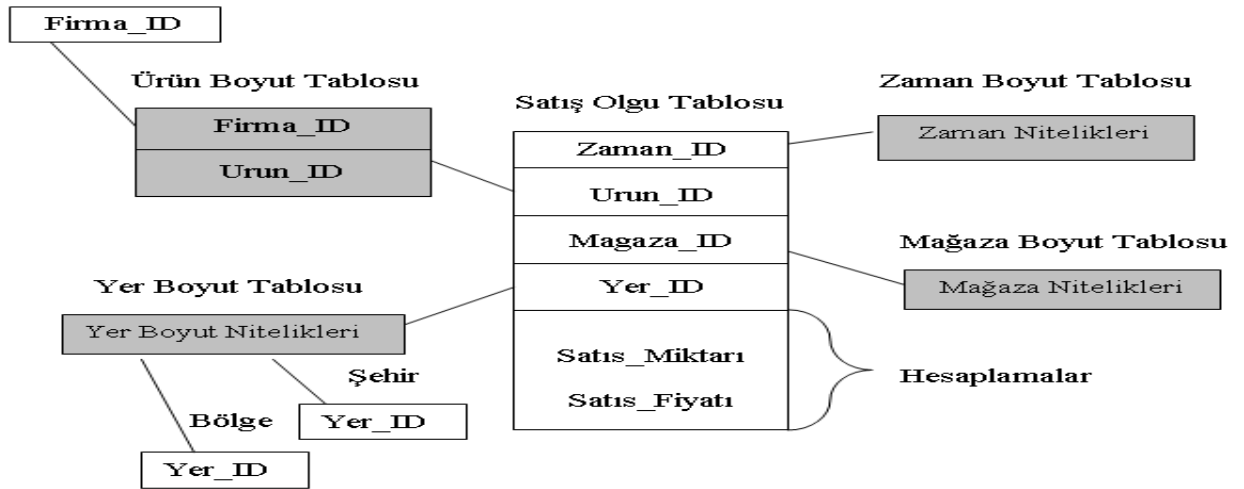
Geleneksel VB modelinde düzenli ve dengeli bir şekilde dağılan varlıklar ve bu varlıklar arasında karmaşık bağıntılar olduğu halde, boyutlu model çok asimetrik bir yapıdadır. Olay tablosu bütün diğer boyut tabloları ile ilişkilendirilmiş olmasına rağmen, olay tablosunu boyut tablolarına bağlayan sadece bir tek ilişki çizgisi (join line) bulunmaktadır.

Kar Tanesi Şema : Kar tanesi şema, bazı boyut tablolarının normalize edildiği, bundan dolayı verinin ek tablolara doğru ileri bölündüğü, yıldız şema modelinin değişik bir biçimidir. Sonuç şema kar tanesine yakın bir şekil oluşturur. [Oracle 10g: Data Warehousing Fundamentals, 2006, pp.3.7]

Kar tanesi ve yıldız şema modelleri arasındaki önemli fark kar tanesi modelinde boyut tablolarının gereksiz fazlalıkları azaltmak için normalize edilmiş formda saklanabilir

olmasıdır. Böyle bir tabloyu yönetmek kolay ve kayıt yerinden tasarruf etmeyi sağlar çünkü büyük bir boyut tablosu, boyutsal yapı olarak sütunlar içerdiğinde devasa hale gelebilir. Bunun yanında yerden kazanç sağlama, olgu tablosunun tipik büyüklüğü ile karşılaştırıldığında önemsizdir. Dahası kar tanesi yapısı, bir sorguyu işletmek için daha çok katılım gerekli olacağından, tarama-gözden geçirme performansının etkinliğini de düşürebilir. Sonuç olarak, sistem performansı ters biçimde etkilenebilir. Bundan dolayı, veri ambarı tasarımında kar tanesi şema, yıldız şema kadar popüler değildir.

Ayrıştırılmış kar tanesi modeli, boyutların sıradüzen yapısını çok güzel bir biçimde görselleştirmektedir. Kar tanesi modelini anlamak boyutları çözümlenecek olan veri modelleyicileri ve veri tabanı tasarımcıları için oldukça kolay olmaktadır. Ancak kar tanesi modelinin daha basit olan yıldız modeline göre biraz daha karışık görünmesi, kullanıcılar açısından daha zor tercih edilmesi sonucunu doğurmaktadır. Ayrıca bu modelin veri saklama gereksinimini azaltması da geliştiricilerin bu modeli elemesine neden olmaktadır. Örneğin, boyutlardan birisi olarak çok büyük bir hesap tablosunun kullanıldığı bir bankacılık uygulamasını ele alalım. Çok tekrarlanan bir metin alanının tablodan çıkarılarak başka bir alt-boyut tablosunda saklanması ve bu sayede yerden kazanılması kolaylıkla akla gelebilir. Kar tanesi modeli yerden kazanç sağlamasına karşın, bu kazanç olay tablosu ile karşılaştırıldığında anlamsız kalmaktadır. Birçok veri tabanı tasarımcısı da modelleme tekniği seçerken, yerden kazanmayı önemli bir ölçüt olarak kabul etmemektedir.



Şekil 2.10. Kar Tanesi Şema

Olgu takımı yıldızı: Karmaşık uygulamalar boyut tablolarını paylaşmak için çoklu olgu tabloları gerektirebilir. Bu çeşit bir şema yıldızların toplamı olarak görülebilir ve bundan dolayı adına galaksi şema veya olgu takımı yıldızı denir.

2.11.1.2.7. Kavram Hiyerarşilerinin Tanıtılması

Kavram hiyerarşisi düşük seviyeli kavramlardan yüksek seviyeli, daha genel kavramlara yapılacak olan anahtarlamaların sıralamasını tanımlar.

Yer boyutu için bir kavram hiyerarşisi düşünün. Yer için şehir değerleri İstanbul, Ankara, İzmir ve Adana yı içermektedir. Bununla birlikte her şehir, ait olduğu bölgeye anahtarlanabilmektedir. Örneğin: İstanbul Marmara ya ve İzmir de Ege ye anahtarlanabilmektedir. Bölgeler, Türkiye gibi sırası ile ait oldukları ülkeye anahtarlanabilmektedir. Yer boyutu için, düşük seviyeli kavramların (Örn: Şehirler) daha yüksek seviyeli-daha genel kavramlara (Örn: Ülkeler) anahtarlanması, bir kavram hiyerarşisi oluşturmaktadır.

Pek çok konsept hiyerarşileri veri tabanı şeması içerisinde gizli kalmaktadır. Örnek olarak, yer boyutunun numara, cadde, şehir, bölge, posta kodu ve ülke nitelikleri ile tanımlandığını düşünün. Bu nitelikler bir toplam sırada, “cadde<şehir<il<ülke” gibi bir kavram hiyerarşisi oluşturacak biçimde, birbiri ile ilişkilidir. Alternatif olarak, bir boyuta ilişkin nitelikler kısmi bir biçimde, kafes oluşturan bir biçimde organize edilebilir. Örnek olarak gün, hafta, ay, çeyrek-üç ay, ve yıl nitelikleri üzerine kurulmuş olan zaman boyutu için kısmi sıralama örneği, “gün<{ay<çeyrek;hafta}<yıl” biçimindedir. Bir haftanın genellikle ardışık olan iki ay arasındaki sınırı geçmesinden dolayı, genelde hafta aya ait düşük bir soyutlama olarak ele alınmaz. Bunun yerine, bir yıl yaklaşık olarak 52 hafta içerdiğinden sıklıkla, hafta yıla ait düşük bir soyutlama olarak değerlendirilir.

Bir veritabanı şeması içerisindeki niteliklerin arasında bulunan toplam veya kısmi kavram hiyerarşisine şema hiyerarşisi denir.

2.12. Veri Ambarı Tasarım ve Geliştirme

Gerekli data kaynaklarının sağlanması ve tutarlı hale getirilen, belirli bir düzene sokulan verinin veri ambarına aktarılması çok önem arz etmektedir.

2.12.1. Veri Ambarı Tasarım Aşamaları

2.12.1.1. Business Model Tanımlama

Veri ambarı gerçekleştirimi için iş süreçlerini tanımlamak için stratejik analizin gerçekleştirilmesi, her bir iş süreci için iş ölçütlerinin ve boyutlarının tanımlanması ve dökümanite edilmesi business model tanımlama kapsamını içerir.

Business Model 3 alan altında incelenebilir [Oracle 10g, a.g.k., pp.3.11]:

2.12.1.1.1. Stratejik Analiz

İş süreçleri seçiminin ve önceliklendirilmesinin yapıldığı business modelin tanımlanması adımının ilk aşamasıdır.

Aşağıdaki stepleri içerir :

- Kurumsal düzeyde en önemli iş süreçlerinin tanımı. Fatura, sipariş, stok, satışlar gibi.
- İş süreçlerini anlamak ve detaylandırmak. İş süreçleri matrisini yaratmak.
- Yatırımsal anlamda geri dönüşü en hızlı olan ve en büyük getiri sağlayacak iş süreçlerini seçmek ve önceliklendirmek.

Tablo 2.1. İş Süreçleri Matrisi örneği

İş Boyutları	İş Süreçleri		
	Satışlar	Dönüşler	Stok
Müşteriler	+	+	
Zaman (Tarih)	+	+	+
Ürünler	+	+	+
Kanallar	+		
Promosyonlar	+	+	

Öncelikler belirlendikten sonra genel strateji belirlenmelidir.

Tüm proje boyunca geçerli olacak strateji, aşağıdaki temel parçalardan oluşur ;

Politika : Data warehouse kullanımına ilişkin çeşitli yönetim kurallarını içerir. Örn: bakım sıklığı, güvenlik yetkileri,arşivleme.

Dönüşüm : Operasyon tarafından transfer edilecek olan dataların tutarlı bir şekilde data warehouse'a aktarılması için dataların temizlenmesi ve dönüştürülmesi.

Data varlığı : Data warehouse'ta tutulacak toplam bilgi kapasitesidir. Bu bilgi gelecekte tekrar tekrar değişim gösterebilir.

Depolama : Veri ambarında tutulacak olan verilerin ulaşılabilir, yönetilebilir, ve esnek olmalıdır. [Bain et al. , 2000, p.128]

2.12.1.1.2. Business Modelin Yaratılması

Veri ambarı gerçekleştirimi için stratejik iş süreçlerinin tanımlandığı bir business model yaratılır.

2.12.1.1.2.1. İş ihtiyaçlarının tanımlanması

Seçilen her bir süreç için iş ihtiyaç analizlerinin tanımlandığı bir model yaratılır.

Aşağıda belirtilen ;

- İş olaylarının tanımlanması ve örneklenerek dökümanite edilmesi,
- Her bir olay için analitik parametrelerin detaylı listesinin yaratılması,
- Gerekli duyarlılık(ayrıntı) tanımlarının yapılması,

- İş tanım ve kurallarının tanımlanması

adımlarının iş sürecinin doğrudan sahibi ve analisti ile bir araya gelerek belirlenmesi gerekir.

Temel veri kaynaklarının doğruluğunun sağlanması için sistematik olarak incelenmesi ve veri denetiminin yapılması çok önemlidir.

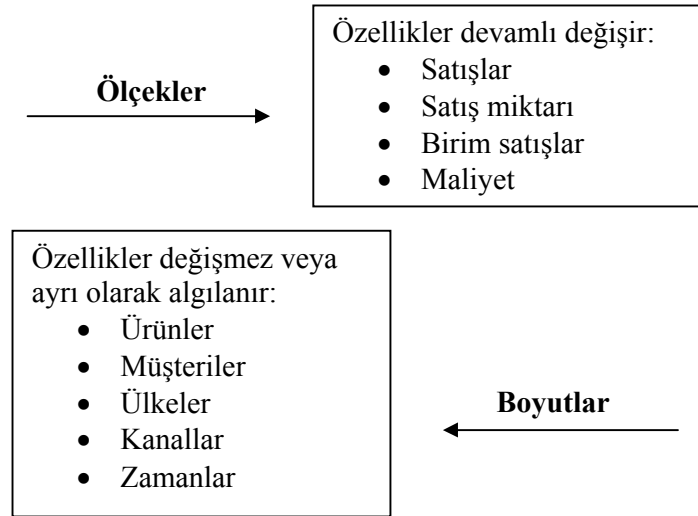
Gereksinimler belgelenmeye başladığı andan itibaren, business modelin de ilk hali şekillenmeye başlar. Gereksinimlerin belirlenmesi tamamlandıkça, business model de buna paralel olarak gelişir.

2.12.1.1.2.2. Boyutlar ve Ölçeklerin Tanımlanması

Veri ambarında izlenmesi zorunlu olan veri elemanları ve bir iş sürecinin başarı metrikleridir. Satış cirosu, toplam maliyet, kar gibi nümerik değerler içerir.

Kurumsal operasyon içerisinde yapılan her iş kurumun işlerini değerlendirmek ya da çözümlenmek amacını taşır. Belirlenen gereksinimlerin, bu çözümlenmenin iki önemli noktasını ortaya koyabilmesi gerekir: neyin çözümlendiği ve çözümlenen konu için değerlendirme ölçütünün ne olduğu. İşte bu bahsedilen değerlendirme ölçütüne ölçek, neyin çözümlendiğine ise boyut diyoruz.

Bir model yaratırken yapılacak ilk iş gereksinimlere göre ölçek ve boyutların ne olduğunun belirlenmesi olacaktır. Bu amaçla, gereksinimleri ifade edebilecek bir soru kümesi oluşturulur ve bu sorular incelenerek, gereksinimleri karşılayacak olan ölçek ve boyutlar belirlenir. Sonraki adımda ise soruları yanıtlayacak veri olup olmadığı ve tasarlanan boyutların bunu karşılayıp karşılamadığı gözden geçirilir.



Şekil 2.11. Boyutlar ve Ölçekler

2.12.1.1.2.3. Duyarlılıkların(Granularity) Belirlenmesi

Boyutlar hakkında daha özel bilgi toplanmak istendiğinde, iş kararları ve analizi için talep edilen konu süreçlerinin detay seviyesini anlamak çok önemlidir. Duyarlılık, konu özet veya detay seviyesinin belirlenmesi ve bilginin o seviyede tutulmasıdır.

Bir olayın duyarlılığı, kaydedildiği ayrıntı düzeyi olarak tanımlanabilir. Eğer verinin etkin bir şekilde çözümlenmesi gerekiyorsa bütün verinin aynı duyarlılık düzeyinde olması gerekir.

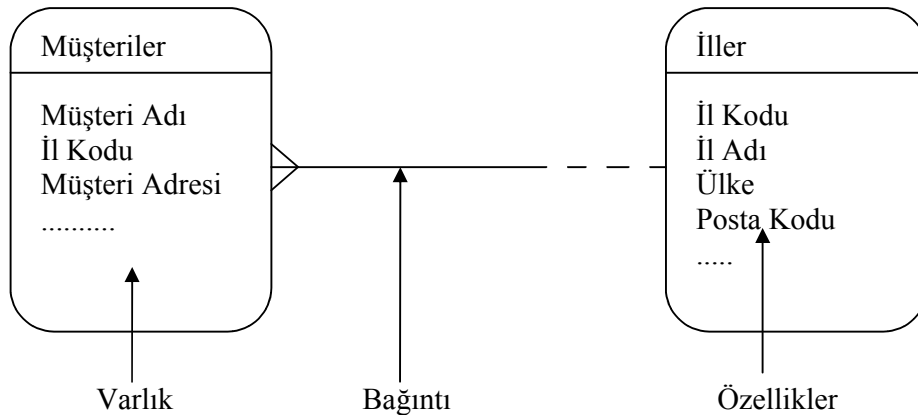
Genel kural, verinin en yüksek duyarlılık düzeyinde saklanmasıdır. Çünkü, verinin hangi ayrıntıda saklanacağına karar verildikten sonra daha fazla ayrıntıya inilmesi mümkün olamaz. Zaman boyutu üzerinde örnekleme yaparsak, veri aydan çeyreğe ve yıla doğru olan bir sıralama ile kümelenebilir. Bu verinin en detay olarak aylık bazda tutulduğunu gösterir. Haftalık veya günlük bazda veriyi görmek isterseniz bunu yapamazsınız.

2.12.1.2. Mantıksal (Logical) Modelin Tanımlanması

Logical design da iş süreçleri, kapsadığı iş konuları ve aralarındaki ilişkiler tanımlanır. Logical design daha çok kavramsaldir. Fiziksel tasarımın soyut şeklidir.

Varlık Bağntı Model (Entity Relationship Model-ERM) organizasyonun mantıksal bilgi gereksinimlerini modellemek için kullanılabilir popöler tekniklerden birisidir. Varlık Bağntı diyagramı ise bir ERM sistemi üzerinde tanımlanmış bilgi gereksinimlerinin görsel bir sunumudur.

Varlıklar, belirli bir konu üzerine bilgilerin tutulduğu tablolar, bağntılar ise bu varlıklar arasındaki ilişkilerin tanımlarıdır.



Şekil 2.12. ER Model

Bağntı derecesi, iki varlık arasındaki ilişkinin derecesini gösterir. Bire bir veya bire çok gibi. Yukarıdaki diyagramda Müşteri ve İl varlıkları arasında bire çok bağntı gösterilmiştir.

2.12.1.3. Boyutlu(Dimensional) Modelin Tanımlanması

Veri modelinin amacı veri gereksinimlerinin açık, net ve özlü bir biçimde ifade edilebilmesidir. Boyutlu model söz konusu olduğunda, gösterimin kullanıcı tarafından anlaşılabilmesi önem kazanmaktadır.

Her bir boyutlu model olay (fact) olarak adlandırılan bir tablo ve boyut (dimension) olarak adlandırılan küçük tablolar kümesinden oluşur.

2.12.1.3.1. Olayların (Facts) Tanımlanması

Boyut kümesi ve bu boyutlarla ilişkili ölçekler olayları oluşturur. Olayları belirlemek üzere boyutlar ve ölçeklerin belirlenmiş gereksinimleri ifade edecek şekilde gruplanarak düzenlenir.

Aynı boyut kümesini tarif eden ölçeklerin hepsine birden tek bir olay yaratılır.

Olay (Fact) tablolar aşağıdaki özelliklere sahiptirler :

- İş amaçlarına uygun sayısal metrikler içerir.
- Büyük volümlü veri tutarlar.
- Çok hızlıdır.
- Temel, türetilmiş ve özet veri içerirler.
- Eklenebilirlik özelliğine sahiptirler.
- Boyut tabloların primary keylerini foreign key olarak içerirler.

2.12.1.3.2. Eklenebilirlik (Additivity)

Eklenebilirlik, ölçeklerin özetlenebilmesidir. Eklenebilirlik, bir olay tablosu üzerinde gerçekleştirilebilecek özetlemeler düşünüldüğünde önem kazanmaktadır.

Ölçekler üçe ayrılır:

- Tam eklenebilir (fully additive),
- Eklenemez (nonadditive),
- Yarı eklenebilir (semiadditive).

Her boyutta eklenebilen boyutlar tam eklenebilir olarak değerlendirilir. Müşteri hesap bakiyesi, birim satışları, nakliye ücretleri olarak örneklendirilebilir.

Yarı eklenebilir ölçekler bazı durumlarda eklenebilen ölçeklerdir. Banka hesap bakiyeleri örnek olarak verilebilir.

Eklenemez ölçekler, hiçbir boyutta eklenemeyen ölçeklerdir. Örnek olarak yüzdesel oranlar verilebilir. Çünkü iki farklı olaydan alınan yüzde oranlarını basitçe birbirine ekleyerek anlamlı bir sonuç çıkarmak zordur.

Ölçekler tam eklenebilir olmalıdır. Aksi takdirde, ölçeklerin daha küçük parçalara ayrıştırılmaları düşünülebilir.

2.12.1.3.3. Zaman Boyutu Eklenmesi

Zaman boyutu veri ambarları için oldukça kritiktir. Verinin zaman dilimleri açısından değerlendirilmesi kurumlar için büyük önem taşır. Kurumlar, planlarını, stratejilerini, hedeflerini genel ve belirli iş konularına göre gruplayarak yaparlar ve belirli zaman dilimlerine göre izleyerek gerçekleştirim oranlarını kontrol ederler. Bunun üzerinden yeni stratejiler geliştirebilirler. Bu bağlamda kurum için zaman boyutunun oluşturulması yerinde olacaktır.

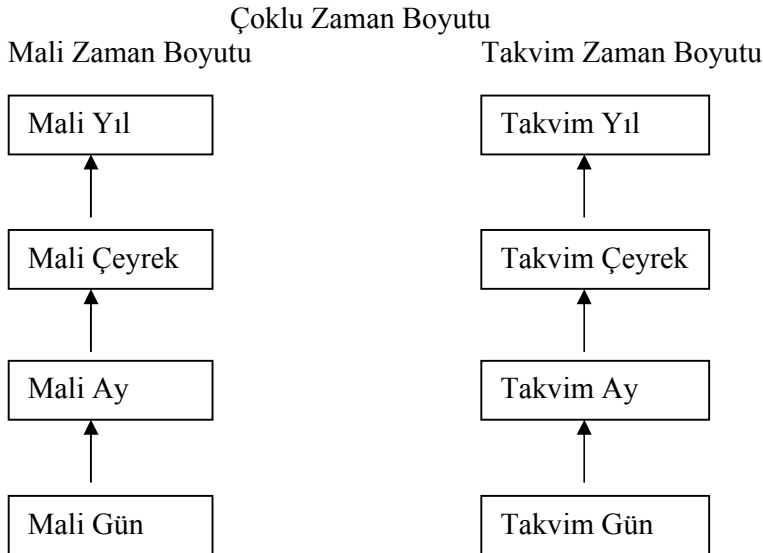
Kurum için oluşturulacak olan zaman diğer boyutlardan bağımsız bir boyut olmalıdır. Birçok kurum için en küçük zaman düzeyi 1 gündür. Gereksinimlerin çözümlenmesiyle birlikte, gün, hafta ve/veya ay bazında raporlama ihtiyacı ortaya çıkabilir.

Zaman boyutu içerisinde gün, hafta, ay gibi niteliklere ek olarak, belirli bir zaman aralığı, 3 ay, bir yıl gibi diğer niteliklerin de değerlendirilmesi iyi olur. Ayrıca tutulan tarihlerin iş günü olup olmadığı, yılın kaçınıcı günü olduğu, yılın kaçınıcı haftası olduğu, yılın başlangıç ve bitiş tarihleri, ayın başlangıç ve bitiş tarihleri, günün adı, bir önceki gün gibi bilgilerin zaman boyutu içinde var olması gereklidir.

Böylece zaman boyutu da belirlendikten sonra, belirlenen gereksinimleri karşılayacak olan boyutlar tamamlanmış olur.

Boyut tabloları hiyerarşik veri içerirler. Tekil veya çoklu olabilirler.

Zaman boyutu üzerinde aşağıdaki şekil 2.13 de çoklu hiyerarşi örneği gösterilmiştir. Sadece mali zaman veya takvim zaman boyutunu göz önüne alırsak tekli zaman boyutu olarak düşünülebilir.



Şekil 2.13. Çoklu Zaman Boyutu

2.12.1.4. Fiziksel Modelin Tanımlanması

İyi yapılmış bir fiziksel model sıklıkla veri ambarının başarı ile başarısızlığı arasındaki farkı gösterir. Fiziksel model tasarımı mantıksal model üzerinde indexlerin eklenmesi, referans bütünlüğü, disk kapasitesi ve diğer özelliklerle inşa edilir.

Fiziksel model tanımlama aşağıdaki işlerin gerçekleştirilmesi ile başarılıdır:

- Boyutsal tasarımın fiziksel modele çevrilmesi.
- Donanım mimarisinin belirlenmesi.
- Tablo ve indeksler için disk stratejisinin tanımlanması.
- Veri tabanı bölümlenmesinin (Database sizing) gerçekleştirilmesi.
- Başlangıç index stratejisinin tanımlanması.
- Partition (bölümleme) stratejisinin tanımlanması.
- Security stratejisinin tanımlanması.

Fiziksel model için başlangıç noktası boyutsal modeldir. Benzer olmalarına rağmen bu iki model arasındaki önemli fark fiziksel modelin, veri tipleri, uzunlukları, veri tabanı büyüklüğü, index ve partition stratejileri gibi fiziki veri tabanı karakteristik özelliklerine sahip olmasıdır.

Veri tabanı objelerinin isim standartlarına sahip olması ve takip edilmesi çok önemlidir. Genelde, mantıksal ve fiziksel isimlemenin mümkün olduğunca tanımlayıcı olmasının yanında özdeş olması da önerilir.

Bu konudaki bazı kurallar aşağıda verilmiştir :

- Tablo ve özelliklerin büyük harflerle isimlendirilmesi.
- Tablo isminin çoğul formda, özelliklerin tekil formda tutulması.
- Objeler isimlerinde kelimeler arasında alt çizgi kullanılması.
- Primary key lerde _PK ekinin kullanılması. URUN_PK gibi.
- Foreign key lerde _FK ekinin kullanılması. KANAL_FK gibi.
- Production keylerde _ID ekinin kullanılması. URUN_ID gibi.

Veri ambarını destekleyecek ortamlar için mevcut teknik mimarinin çerçevesi, ana hatları ve planı sağlanmalıdır. Mevcut ortamların varolan kapasite planları göz önünde bulundurulmalıdır. Veri ambarının kapasite planı, teknik mimarisi ve geri dönüş stratejisi hazırlanmalıdır.

Donanım mimarisi ölçeklenebilir, yönetilebilir, kullanılabilir, genişleyebilir, esnek ve entegre özelliklerine sahip olmalıdır.

Veri ambarını oluşturacak tablo sayısı, tablo kayıt boyu, kayıt sayısı ve indexler gözönüne alınarak veri ambarı büyüklüğü hesaplanır. Mutlaka tarihsel olarak tutulacak veri ve kurumun büyüme stratejisi gözönüne alınarak gelecek veri tahminleri veri ambarı büyüklüğünü hesaplarken gözönünde bulundurulmalıdır. Bu donanım mimarisi için önemli adımlardan birisidir.

2.13. Kaynakların Tanımlanması

Modeli yükleme aşamasına geldiğimizde öncelikle kullanılacak olan veri kaynakları belirlenir. Kaynaklar daha sonra veri ambarı modeline eşlenir. Bu eşleme her olay, ölçek, boyut ve boyut alanları için yapılır. Her boyut ve boyut alanları için, sadece kaynak varlıklar (ilişkisel tablolar, kütükler, ...) belgelendirilmelidir. Dönüştürme ve türetme algoritmaları da belirlenmelidir.

Boyut alanları düzeyinde bu işlem, veri türü dönüşümü, kaynak alanlarını birleştirme ve bölme algoritmalarını, yapılması gereken hesaplamaları, alan dönüşümleri ve kaynak seçimini içerir.

Alan dönüştürme, kaynak sistemdeki alanı, hedefteki yeni değer kümesine değiştirme işlemidir. Kullanılan operasyonel sistem medeniyet durumu için, 1=bekar, 2=evli, 3=boşanmış kodlamasını kullanıyor olsun. Bu kodlama hedef sistemde bekar, evli ve boşanmış olarak değiştirilebilir.

Hedef alan farklı kaynaklardaki alanlardan da yüklenebilir.

Kaynak tanımlama, modelin değişmesine neden olabilir. Kaynağı olmayan hedef tespit edildiğinde bu durum gerçekleşir. Bu durum da model yeniden gözden geçirilir.

2.14. Veri Kalitesi

Veri kalitesi, veri temizliği, verilerde standartlaşmaya gidilmesi, veri kümelerinin genişletilmesi, çift ürün ve müşteri kayıtlarının tekilleştirilmesi işlemlerinin bütününe verilen isimdir.

Veri kalitesi metodolojisinde, bir kurumun veri kalitesini arttırmaya başlayacağı ilk noktayı veri keşfi (data profiling) olarak nitelendirilmektedir. Veri Keşfi sırasında verilerin sayısallaştırılması ve analiz edilmesi mümkündür. Bu sayede örneğin bir tablo alanı içindeki verilerin kendini ne kadar tekrar ettiği, ne kadarının boş, ne kadarının dolu olduğu, ya da iki tablo arasındaki tutarsız bilgilerin oranı gibi bilgilere sayısal olarak erişmek ve gerektiğinde detaya inerek bunların hangi kayıtlardan geldiğini görmek mümkündür.

Veri temizleme (Data Cleansing) kapsamında verilerin ayrıştırılması (parse), standartlaştırılması, yanlış bilgilerin düzeltilmesi, veri kümelerinin genişletilmesi

(enhancement), müşteri kayıtlarının birbiriyle eşleştirilmesi ve tekilleştirilmesi temel olarak veri kalitesini oluşturmanın veri keşfinden sonraki adımlarıdır.

2.15. Veri Ambarı Üretim Süreci (ETL - Extract, Transform, Load)

Veri ambarı üretim süreci, farklı operasyonel sistemlerdeki verilerin çekilerek transformasyona tabi tutulduktan sonra bilgilerin veri ambarına ya da veri pazarına yüklenmesi anlamına gelmektedir. Gerçek zamanlı veri besleme de (Real-time data feed) mümkündür.

2.15.1. Veri Yakalama (Capture)

Veri yakalama işlemi, operasyonel sistemler ve diğer dış kaynaklardaki verilerin toplanması işlemidir. Yakalama işlemindeki veri kaynağı, dosya biçimlerini, ilişkisel ve ilişkisel olmayan veri tabanı yönetim sistemlerini içermektedir. Verinin ne türde bir dosyadan yakalandığı, veri yakalamada kullanılan tekniğe bağlıdır. Bu teknikler ;

- Kaynaktan veri çekme (source data extraction),
- VTYS günlük yakalama (DBMS log capture),
- Tetiklenmiş yakalama (triggered capture),
- Uygulama yardımıyla yakalama (application-assisted capture),
- Zaman-damgasi-tabanlı yakalama (time-stamp-based capture)
- Karşılaştırmalı yakalama (comparison capture)

teknikleridir.

Kaynaktan veri çekme, kaynak verinin belli bir zamandaki statik görüntüsünü verir. Bu teknik, sürekli bir tariheye ihtiyaç duyulmadığında kullanılabilir. Bu işlem, veri bloğunun tamamı alındığında ve depodaki mevcut veri aynı kaynaktan ancak farklı zamandaki yenisiyle değiştirildiğinde iyi sonuçlar vermektedir. Okul bu tekniği kullanımı için iyi bir örnek olabilir. Okulun veri ambarı yıllık olarak güncellenir ve önceki yıla ait eski veri, yıl sonunda arşive aktarılır.

VTYS Log Yakalama, verinin VTYS günlük sisteminden yakalanmasını sağlar. Bu teknik, sadece ilgilenilen verinin yakalanabilmesini sağlamak üzere, günlük kayıtların nasıl bir biçimde tutulduğunun bilinmesini gerektirir. Ayrıca, log kayıtlarının formatının kolayca anlaşılabilmesi için net olmalıdır.

Tetiklemeli yakalama, tetikleyiciler ve saklı prosedürler, ekleme, güncelleme ve silme senaryoları gibi önceden tanımlı olaylar gerçekleştiğinde çalışmak üzere hazırlanmakta ve pek çok veritabanı yönetim sistemi tarafından desteklenmektedir. Tetikleyici, belirlenen olayın gerçekleştiğini farkeder ve prosedürü aktif hale getirir. Prosedürün kontrol ve bakımı,

kullanıcının sorumluluğundadır. Tetikleyici, veritabanı yönetim sisteminin değil prosedürü yazanın denetiminde olduğundan, değişikliklere de açıktır.

Uygulama destekli yakalama, operasyonel kaynaklardan veri toplamak için program yazımını içeren tekniktir. Bu teknik, test ve bakım sorumluluklarını içerecek şekilde, tamamen programcının denetimi altındadır.

DBMS log yakalama, tetiklemeli yakalama ve uygulama destekli yakalama metotlarının, sürekli tarihsel modelle çalışmak için artışı bir kayıt yapısı sağladığından bahsetmek mümkündür .

Zaman damgası tabanlı yakalama, kaydın son yakalamadan sonra değişip değişmediğini belirlemek üzere kullanılan bir zaman damgasına dayanan bir tekniktir. Eğer bir kayıt değişmişse ya da yeni bir kayıt eklenmişse, bu kayıt işlenmek üzere başka bir dosya veya tabloya aktarılır.

Dosya karşılaştırma, veri kaynağının belirli bir zamanda alınmış enstantenesi, bir dosyaya kaydedilir ve sonra daha önceki bir halini içeren dosya ile karşılaştırılır. Belirlenen herhangi bir değişiklik ya da ekleme, daha sonra işlenmek üzere ayrı dosyalara koyulur.

2.15.2. Veri Dönüştürme (Transform)

Dönüştürme işlemi, yakalanan kaynak veriyi veri ambarına yüklenebileceği uygun biçim ve yapıya dönüştürür. Dönüştürme işlemi, kaynak verideki bozuklukları ortadan kaldırarak hedef veri ambarı için yüksek kaliteli verinin elde edilmesini sağlar. Verinin dönüştürülmesi, kayıt ya da nitelik düzeyinde gerçekleştirilebilir. Temel dönüştürme teknikleri, yapısal dönüştürme (structural transformation), içerik dönüştürme (content transformation) ve işlevsel dönüştürme (functional transformation) teknikleridir.

Yapısal dönüştürme, kaynak kayıtların yapısını hedef veri tabanına göre değiştirme işlemidir. Bu teknikte veri kayıt düzeyinde dönüştürülür.

İçerik dönüştürme, kayıtlardaki verilerin değerlerini değiştirme işlemidir. Bu işlem, veriyi nitelik düzeyinde dönüştürmektedir. İçerik dönüştürme, algoritmalar ya da veri dönüştürme tabloları kullanılarak gerçekleştirilebilir.

İşlevsel dönüştürme, kaynak kayıtlardaki verilere dayanarak, hedef kayıtlarda yeni veri değerleri yaratılması işlemidir. Bu teknikte de veri nitelik düzeyinde dönüştürülür.

2.15.3. Uygulamaya Koyma (Apply)

Uygulamaya koyma işlemi, dönüştürme işlemi ile yaratılan dosyaları işleyerek, bunları ilgili veri ambarı ya da veri pazarına yüklemektir.

Bu işlem için 4 temel teknik kullanılır :

- Yükleme (load),

- Ekleme (append),
- Yapıcı birleştirme (constructive merge),
- Yıkıcı birleştirme (destructive merge).

Yükleme (Load), hedef veri ambarı tablolarını, dönüştürme işleminde yaratılan tablolara değiştirir. Eğer veri ambarında bu tablolar yoksa, tablolar yaratılır.

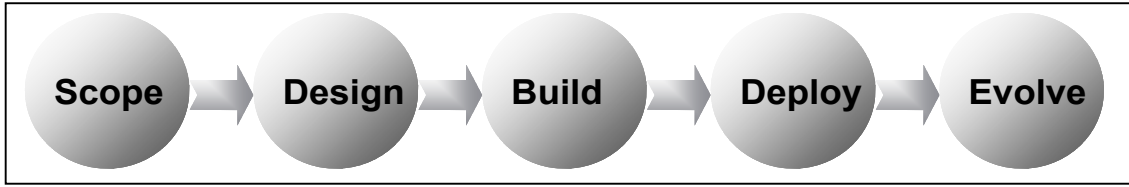
Ekleme (Append), varolan tabloların sonuna dönüştürme sonucunda ortaya çıkan tablodaki yeni verileri ekler.

Yapıcı birleştirme (Constructive merge), var olan hedef tablolara yeni verileri ekleyerek durumu değişen kayıttaki bir zaman değerini günceller.

Yıkıcı birleştirme (Destructive merge), yeni verileri var olan kayıtların üzerine yazar.

2.16. Veri Ambarı Proje Yönetim Aşamaları


Veri Ambarı projeleri genelde aşağıda gösterilen 5 adımda gerçekleştirilir.



Şekil 2.14. Veri Ambarı Proje Yönetim Aşamaları

Aşağıdaki tabloda , bu adımlar hakkında kısa açıklamalar ve bu adımların her birinin sonunda elde edilecek üretimlerin neler olacağı belirtilmiştir.

Tablo 2.2. Veri Ambarı Proje Yönetim Adımları ve Detaylar

Proje Adımı	Açıklama	İlgili Adım Sonunda Yapılan Teslimatlar
	Scope aşamasının amacı, proje kapsamının ve şartlarının tanımlanması, kurumun mevcut durumunun analiz edilmesi, iş ihtiyaçlarının detaylarıyla birlikte anlaşılması, performans göstergelerinin belirlenmesi ve tanımlanması ve önemli iş akışlarının zayıf noktalarıyla birlikte ortaya çıkarılması olarak özetlenebilir. Bu fazın sonunda tüm fonksiyonel ve teknik ihtiyaçlar belirlenmiş ve tanımlanmış çözümün tam bir resmi ortaya konulmuş olacaktır. Scope aşamasında, iş	a. Proje Tanım Dökümanı b. Revize Proje Planı

	birimlerindeki yöneticiler/analistlerle gerekli toplantılar yapılacak ihtiyaçların ne olduğu net olarak anlaşılacaktır.	
Design	Scope aşamasında tanımlanmış ihtiyaçlar daha detaylı incelenecek ve teknik altyapının oluşturulabilmesi için gerekli veri analizi yapılacaktır. Oluşturulacak veri ambarına ait detay seviyesi (Granülarite) bu aşamada finalize edilecektir. Raporlar için gereken veri modeli, ETL ve mimari bu aşamada netleştirilecektir.	<ul style="list-style-type: none"> a. Veriye ait tanımlar b. Mantıksal Veri Modeli Spesifikasyonları c. ETL Tasarım Spesifikasyonları d. Rapor Spesifikasyonları e. Dashboard Spesifikasyonları f. Mimari Spesifikasyonlar
Build	Bu aşamada, Design fazında belirlenmiş spesifikasyonlara uygun ETL modüllerinin hazırlanması, verilerin operasyonel sistemlerden çekilerek hedef veri tabanına aktarılması, universe tasarımının gerçekleştirilmesi ve raporların oluşturulması işlemleri tamamlanacaktır. Bu faz, gerçek anlamda uygulama geliştirmenin yapıldığı fazdır. Build aşaması, verinin kaynak sistemlerden çekilmesi, dönüştürülmesi, yüklenmesi aşamalarında önceliklendirilmiş ihtiyaçları karşılayacak düzeye gelene kadar, döngüsel bir karakter sergileyecektir.	<ul style="list-style-type: none"> a. Fiziksel Veri Modeli b. ETL Scriptleri c. Kapsam dahilinde belirlenmiş Universe'ler d. Kapsam dahilinde üretilecek raporlar e. Dashboard ekranları f. Veri Ambarında test datanın hazırlanması
Deploy	Bu aşama, sistemdeki tüm bileşenlerin bir araya getirilerek production ortamına sunulduğu ve testlerin yapıldığı aşamadır. Bu aşamada önce initial load tamamlanacak ve ardından tüm fonksiyonel ve entegrasyon testleri gerçekleştirilecektir.	<ul style="list-style-type: none"> a. Initial Load bitmiş olacak b. Sistem Testleri tamamlanacak
Evolve	Bu aşamada sistem kısıtlı bir kullanıcı kitlesi için production ortamında sunulacaktır. Bu kullanıcılar öncelikle eğitilecek ve hemen ardından kabul testlerine başlayacaktır. Proje Tanım dökümanında belirtilen maddelerden sapmalar ya da yapılması gereken son düzeltmeler bu aşamada gerçekleştirilecektir.	<ul style="list-style-type: none"> a. Örnek bir kullanıcı grubunun eğitimleri verilecek, b. Kabul testleri tamamlanacak c. Sistem tüm kullanıcılara açılacak

III. BÖLÜM : VERİ MADENCİLİĞİ

3.1. Veri Madenciliği (Data Mining)

Veri kendi başına değersizdir. İstedığımız amacımız doğrultusunda bilgidir. Bilgi bir amaca yönelik işlenmiş veridir. Veriyi bilgiye çevirmeye veri analizi denir. Bilgi de bir soruya yanıt vermek için veriden çıkardığımız olarak tanımlanabilir. Veri sadece sayılar veya harfler değildir; veri, sayı ve harfler ve onların anlamıdır.

Teknolojinin büyük hızla gelişmesi sonucu durmadan büyüyen ve işlenmediği sürece değersiz gibi görünen veri yığınları oluşmaktadır. Bu veri yığınlarını, içlerinde altın madenleri bulunan dağlara benzetmek mümkündür. Bu madenlere ulaşmak için kullanılan yöntem ise, temelinde istatistik uygulamaları yatan “VERİ MADENCİLİĞİDİR”.

Veri madenciliği en basit tanımı ile çok büyük miktardaki ham veriler içinden amaca uygun modellerin ortaya çıkarılması işlemidir. Başka bir tabirle karmaşık ve düzensiz veriler içindeki modelleri (ilişkiler, örüntüler, değişiklikler, sapma ve eğilimler, belirli yapılar gibi ilginç bilgileri) ortaya çıkarıp bunları karar verme ve eylem planını gerçekleştirmek için kullanma sürecidir.

Veri Tabanlarında Bilgi Keşfi (Knowledge Discovery in Databases) uygulamaları ile birlikte faaliyet alanına yönelik karar destek mekanizmaları için gerekli ön bilgileri temin etmek için kullanılır. [Fayyad et al., 1996a]

Geleceğin, en azından yakın geleceğin, geçmişten çok fazla farklı olmayacağını varsayarsak geçmiş veriden çıkarılmış olan kurallar gelecekte de geçerli olacak ve ilerisi için doğru tahmin yapmamızı sağlayacaktır.

Verileri kaydetmek, yalnızca maden yataklarının yerlerini tespit etmektir. Bu veriler operasyonel amaçlarla kullanılırsa bir çöp olmaktan ileriye gidemezler. Veri çöplüğünden kurtulup değerli bir veri madenine sahip olmak için elimizdeki bu bilgileri değerlendirmemiz şarttır.

Bir altın madeninde kazı yapacak olsanız, altını çıkarmak için ekonomik değeri olmayan bir sürü madde içerisinden altını çıkartmanız, ve çıkardığınız bu altını işleyerek ona değer katmanız gerecektir. Bir veri madeninde değerli bilgiler ararken de durum pek farklı değil, gerekli bilgiyi çıkarıp, daha sonrasında da bu bilgiyi işlemek için stratejiler uygulamazsak sonuca ulaşmamız mümkün değildir.

Veri madenciliği tanımlarında öne çıkan noktalar şunlardır:

- Büyük ve karmaşık verilerle çalışır.
- Her türlü veriyi kullanarak çözümler üretebilir.

- İstatistik, yapay zeka, makine öğrenmesi, Veri tabanlarında bilgi keşfi, bilgisayar
- bilimi, yapı tanıma vb. gibi disiplinlerden faydalanır.
- Daha önceden bilinmeyen, doğrulanabilir, etkinleştirilebilir enformasyon arar.
- Otomatik veya yarı otomatik olarak çalışan çözüm araçları kullanır.
- Birçok endüstride kullanılmaktadır.
- Sorunlara göre değişen çözüm araçları vardır.

3.2. Veri Madenciliği İşlevi

Veri madenciliğine işlevleri açısından bakılacak olursa, veri madenciliği aktiviteleri 3 sınıf altında toplanmıştır :

- Keşif (discovery),
- Tahmini modelleme (predictive modeling),
- Adli analiz (forensic analysis).

Keşif , ne olabileceği konusunda önceden belirlenmiş bir fikir ya da hipotez olmadan, veri tabanı içerisinde gizli desenleri arama işlemidir. Geniş veri tabanlarında kullanıcının pratik olarak aklına gelmeyecek ve bulmak için gerekli doğru soruları bile düşünemeyeceği birçok gizli desen olabilir. Buradaki asıl amaç, bulunacak desenlerin zenginliği ve bunlardan çıkarılacak bilginin kalitesidir.

Tahmini modelleme, veri tabanından çıkarılan desenler, geleceği tahmin için kullanılır. Bu model, kullanıcının bazı alan bilgilerini bilmese bile kayıt etmesine izin verir. Sistem, bu boşlukları, önceki kayıtlara bakarak tahmin yoluyla doldurur. Keşif, verideki desenleri bulmaya yönelikken, tahmini modelleme, bu desenleri yeni veri nesnelere bulmak için uygular.

Adli analiz, normal olmayan ya da sıra dışı veri elemanlarını bulmak için, çıkarılmış desenleri uygulama işlemidir. Sıra dışı olanı bulmak için ilk önce sıradan kısmı tespit etmek gerekir.

3.3. Veri Madenciliğinin Karar Verici için Yararları

Veri madenciliğinin karar vericiler için sağladığı yararlar aşağıda sıralanmıştır.

- Mevcut müşterilerin karar verici tarafından daha iyi tanınmasını sağlayabilir.
- Özellikle finans sektöründe mevcut müşterileri bölümlere ayırıp, kredi risk davranış modelleri oluşturarak, yeni başvuruda bulunan müşterilere karşı riskin minimize edilmesini sağlayabilir

- Mevcut müşterilerin ödeme performansları incelenip kötü ödeme performansı gösteren müşterilerin ortak özellikleri belirlenerek, benzer özelliklere sahip tüm müşteriler için yeni risk yönetim politikaları oluşturulabilir.
- En iyi müşteriler veya müşteri bölümlerinin bulunmasında kullanılabilir. Bulunan bu iyi müşteri bölümlerine yönelik yeni pazarlama stratejileri oluşturulabilir.
- Kuruluşlar tarafından düzenlenecek çeşitli kampanyalarda mevcut müşteri kitlesinin seçimi ve bu müşterilerin davranış özelliklerine yönelik kampanya şartlarının oluşturulması sağlanabilir.
- Bankacılık faaliyetlerinde, küçük işletmelere yönelik olarak makine ve ekipman satışı yapan dağıtıcı firmalarla ortak hareket ederek oluşturulacak satış paketleri ile pazarlama stratejileri geliştirilebilir.
- Mevcut müşteriler üzerinde firma ürünlerinin çapraz satış kapasitesinin artırılması sağlanabilir.
- Veri madenciliği ile mevcut müşteriyi tanıyarak kuruluşların müşteri ilişkileri yönetimlerinde düzenleme ve geliştirmeler yapılabilir. Bu sayede kuruluşun müşterilerini daha iyi tanıyarak müşteri gibi düşünme kapasitelerinin artırılması sağlanabilir.
- Günümüzde var olan yoğun rekabet ortamında kuruluşların hızlı ve kendisi için en doğru kararı almalarını sağlayabilir.
- Kuruluşlar veri analizi ile müşterilerini kişiselleştirilmiş ürün ve hizmetler hakkında bilgilendirebilirler.
- Veri madenciliği ile kuruluşların müşteriyle bütünleşmiş satış politikaları oluşturması sağlanabilir.

3.4. Veri Tabanlarında Bilgi Keşfi (VBK) Süreci (Knowledge Discovery in Databases)

Veri Madenciliği gözlemsel veritabanlarındaki veriler arasındaki ilişkilerin ve örüntülerin keşfinde kullanılmaktadır.

VM, “veri tabanlarında bilgi keşfi (VTBK)” sürecinin ayrılmış bir adımı olarak ele alınabilir (Şekil 3.1). Bu VTBK süreci verilerden kullanışlı tüm bilgi keşfi süreçlerini kapsamaktadır.

VTBK süreçlerinde ek olarak veri hazırlama, veri seçimi, veri temizleme, daha önce elde olan uygun bilgilerle birleştirme ve madencilik sonuçlarının uygun yorumu, verilerden faydalı bilgi çıkarımını sağlamak için gereklidir.

Veri tabanında bilgi keşfi, verilerin doğru, faydalı ve anlaşılır modeller ve kalıplar elde etmede kullanılan özel bir süreçtir.

Bu tanımdaki “özel bir süreç” ifadesinden de anlaşılacağı gibi Veri Tabanlarında Bilgi Keşfi Süreci(VTBK), karmaşık bir işlemdir. “Model elde etmek” ile kastedilen ise, verilere model uydurmak, verilerden yapı ortaya çıkarmak veya genel bir ifadeyle, veri kümesine yüksek dereceli açıklama getirmektir

“Süreç” ise, VTBK’nin birçok adımdan ve çeşitli iterasyonlardan oluştuğunu göstermektedir. Bilgi keşfinin belirli bir güven düzeyinde, yeni veriler için geçerli olması, süreç sonunda elde edilen bilginin ise, iş konusunda çeşitli avantajlara olanak sağlayacak şekilde faydalı ve anlaşılır olması gerekmektedir. [Fayyad et al. , 1996b]

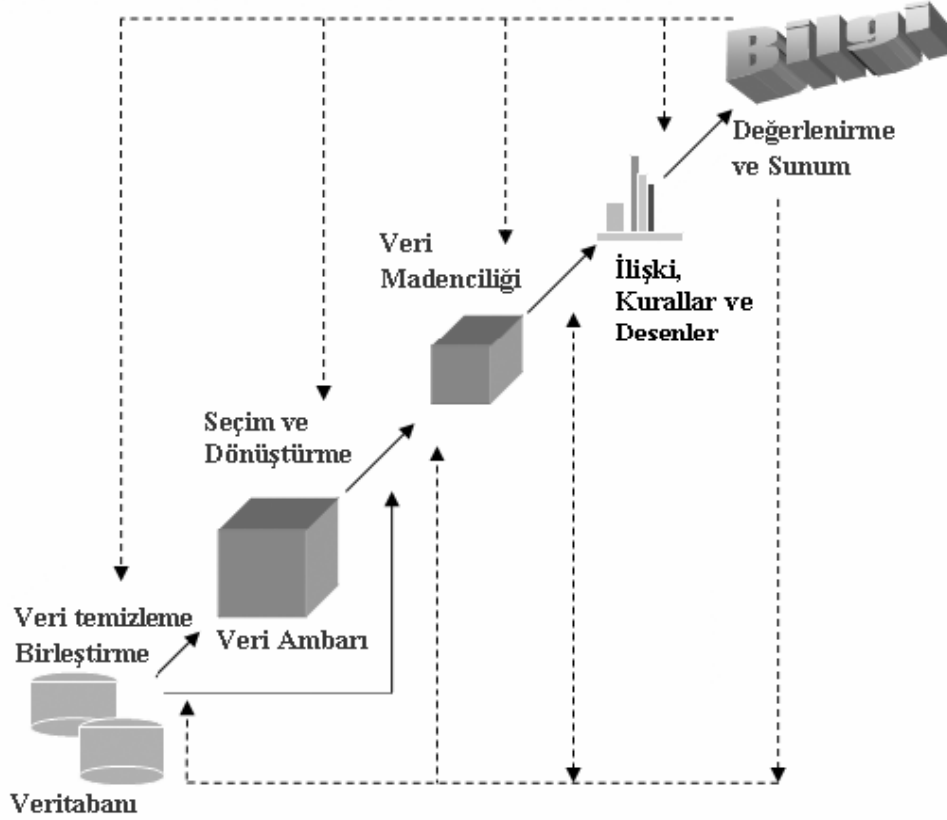
1. Veri Seçimi : Bu adım veri kümelerinden sorguya uygun verilerin seçilmesidir. Elde edilen verilere örneklem kümesi denmektedir.

2. Veri Temizleme ve Ön işleme : Örneklem kümesi elde edildikten sonra, örneklem kümesinde yer alan hatalı tutanakların çıkarıldığı ve eksik nitelik değerlerinin değiştirildiği aşamadır. Bu aşama seçilen veri madenciliği sorgusunun çalışma zamanını iyileştirir.

3. Veri Madenciliği : veri temizleme ve ön işlemde geçen örneklem kümesine VM sorgusunun uygulanmasıdır. Örnek VM sorguları: kümeleme, sınıflandırma, ilişkilendirme, vb. sorgulardır.

4. Yorumlama : VM sorgularından ortaya çıkan sonuçların yorumlanma kesimidir. Burada geçerlilik, yenilik, yararlılık ve basitlik açılarından üretilen sonuçlar yorumlanır.

Veri tabanında bilgi keşfi süreci basamakları basamaklar Şekil 3.1’de ifade edilmiştir.



Şekil 3.1. Bilgi keşfi sürecinde veri madenciliği (Data Mining in Knowledge Discovery)

3.5. Veri Madenciliğindeki Problemler

Veri madenciliği girdi olarak ham veriyi sağlamak üzere veri tabanlarına dayanır. Bu da veri tabanlarının dinamik, eksiksiz, yeterli sayıda ve net veri içermemesi durumunda sorunlar doğurur. Sınıflandırmak gerekirse başlıca sorunlar şunlardır:

Sınırlı Bilgi : Veri tabanları genel olarak basit öğrenme işlerini sağlayan özellik veya nitelikleri sunmak gibi amaçlar için hazırlanmışlardır. Bu yüzden, öğrenme görevini kolaylaştıracak bazı özellikler bulunmayabilir.

Veri tabanı boyutu : Veri tabanı boyutları inanılmaz bir hızla artmaktadır. Veri tabanı algoritması çok sayıda küçük örnekleme ele alabilecek biçimde geliştirilmiştir. Aynı algoritmaların yüzlerce kat büyük örneklerde kullanılabilmesi için çok dikkat gerekmektedir. Örneklemin büyük olması, tahminlerin doğruluğu açısından bir avantaj olsa da dikkatsizlikten kaynaklanacak hatalar göz ardı edilemez.

Aykırı veri : Veri girişi veya veri toplanması sırasında oluşan sistem dışı hatalara gürültü adı verilir. Verilerde ne kadar çok gürültü varsa o derece güvenilir sonuçlara ulaşmak zorlaşacaktır. Bu gürültüler geleceğe dair tahminlerin doğruluğunun azalmasına neden olur. Gürültülü verilerden kurtulmak için yanlış, çok fazla ya da çok küçük araştırmalara dair

tutarsız bilgiler yerine anlamlı, özümsemiş bilgiler kullanılmalıdır. Gürültülü verilerin teşhis edilmesi amacıyla histogram, kümeleme analizi ve regresyon kullanılır.

Eksik veri : Veri kümesinin büyüklüğünden ya da doğasından kaynaklanmaktadır. Eksik veriler olduğunda yapılması gerekenler şunlardır:

- Eksik veri içeren kayıt veya kayıtlar çıkarılabilir.
- Değişkenin ortalaması eksik verilerin yerine kullanılabilir.
- Var olan verilere dayalı olarak en uygun değer kullanılabilir.

Eksik veriler, yapılacak olan istatistiksel analizlerde önemli problemler yaratmaktadır. Çünkü istatistiksel analizler ve bu analizlerin yapılmasına olanak veren ilgili paket programlar, verilerin tümünün var olduğu durumlar için geliştirilmiştir. Bu analizler, eksik veri içeren veri setlerine uygulandıklarında istatistiklerin geçerliliğini düşürmektedir.

Null değerler : Eğer VT'de bir nitelik değeri NULL ise o nitelik bilinmeyen ve uygulanamaz bir değere sahiptir. VT'de birincil anahtar haricindeki herhangi bir niteliğin özelliği NOT NULL (NULL olamaz) şeklinde tanımlanmadığı sürece bu niteliğin değeri NULL olabilir.

Veri kümelerinde var olan NULL değerleri için çeşitli çözümler söz konusudur [Quinlan,1986] :

- NULL değerli kayıtlar tamamıyla ihmal edilebilir,
- NULL değerli kayıtlardaki NULL değerleri olası bir değerle günlenebilir. Bu günleme için çeşitli yöntemler söz konusudur:
 - o NULL değeri yerine o nitelikteki en fazla frekansa sahip bir değer veya ortalama bir değer konulabilir,
 - o NULL değeri yerine varsayılan bir değer konulabilir,
 - o NULL değerinin bulunduğu kaydın diğer özelliklerine göre, NULL değerinin kendine en yakın değerle günlenmesi sağlanabilir, vb.

Artık veri, problemde istenilen sonucu elde etmek için kullanılan örneklem kümesindeki gereksiz niteliklerdir. Artık nitelikleri elemek için geliştirilmiş algoritmalar, özellik seçimi olarak adlandırılır. Özellik seçimi arama uzayını küçültür ve sınıflama işleminin kalitesini de artırır [Deogun et al., 1995; Kira and Rendeli, 1992; Almuallim and Dietterich, 1991; Pawlak, 1986].

Dinamik veri, içeriği sürekli değişen veri tabanlarıdır. Bunlara örnek kurumsal çevrim-içi veri tabanları gösterilebilir. Bir veri tabanındaki içeriğin sürekli değişmesi VM

uygulamalarının uygulanabilmesini önemli ölçüde zorlaştırıcı sorunlar doğurmaktadır. Bu sorunlardan bazıları şunlardır:

- Ortaya çıkan VM örüntülerinin sürekli değişim halinde olan verilerden hangisini ifade ettiğinin tespitinin zorluğu ve bu üretilen sonuçların zaman içinde eski üretilen sonuçlardan farkının tespiti ve gereken yerlerin güncellenme zorluğu,
- VM algoritmalarının çalışabilmesi için verilerin üzerine okuma kilidi konulması gerektiğinde, bu verilerin başka uygulamalar tarafından değişime açık olmaması,
- VM algoritmalarının ve çevrim-içi VT uygulamalarının aynı anda uygulanmasından kaynaklanan ciddi performans düşüşlerinin olması, vb.

3.6. Veri Madenciliğinde Kullanılan Yöntemler

Veri madenciliği sürecinin çeşitli aşamalarında kullanılan teknikler, istatistiksel yöntemler, bellek tabanlı yöntemler, genetik algoritmalar, yapay sinir ağları ve karar ağaçları olarak sıralanabilir.

3.6.1. İstatistiksel Yöntemler

Veri madenciliği çalışması esas olarak bir istatistik uygulamasıdır. Verilen bir örnek kümesine bir kestirici oturtmayı amaçlar. İstatistik literatüründe son yıllarda bu amaç için değişik teknikler önerilmiştir. Bu teknikler istatistik literatüründe çok boyutlu analiz başlığı altında toplanır ve genelde verinin parametrik bir modelden (çoğunlukla çok boyutlu bir Gauss dağılımından) geldiğini varsayar. Bu varsayım altında ;

- Sınıflandırma : Sınıflandırma, yeni bir nesnenin niteliklerini inceleme ve bu nesneyi önceden tanımlanmış bir sınıfa atamaktır. Burada önemli olan, her bir sınıfın özelliklerinin önceden net bir şekilde belirlenmiş olmasıdır. Sınıflandırmaya örnek olarak kredi kartı başvurularını düşük, orta ve yüksek risk grubu olarak ayırmak gösterilebilir.
- Ayırma Analizi : Ayırma analizi, iki veya daha fazla sayıdaki grubun ayırımı ile ilgilenen çok değişkenli ilgi analizidir. Amaçları arasında, analiz öncesi tanımlanmış iki veya daha fazla sayıda grubun ortalama nitelikleri arasında önemli farkların olup olmadığının test edilmesi, gruplar arasındaki farka her bir değişkenin katkısının saptanması ve grup içi değişime oranla gruplar arasındaki ayırımı maksimize eden tahmin değişkenleri kombinasyonunun belirlenmesi sayılabilir. Örneğin, bira içenleri, bira içmeyenlerden ayırt etmenin bir pazarlama sorunu olduğu kabul edilirse, büyük

bir bira üreticisinin yaptığı araştırma ayırma analizine örnek olarak gösterilebilir. Bu nedenle, tesadüfi olarak seçilen 500 kişilik bir tüketici bölümünü örnek olarak alınmış ve bu kişilerin bira içip içmedikleri, cinsiyetleri ve sporla ilgilenme dereceleri saptanmıştır. Cinsiyet ve sporla ilgilenmenin tahmin değişkenleri olarak kullanılmalarının nedeni, daha önceki çalışmaların bu değişkenlerle bira içme arasında kuvvetli bir ilginin olduğunu göstermiş olmasıdır. Ayırma analizi sonuçlarının test edilme olanağının bulunması sonuçların geçerliliğini ve güvenilirliğini ve dolayısıyla analizin gücünü artıran önemli bir etmendir.

- Regresyon : Bir ya da daha çok değişkenin başka değişkenler cinsinden tahmin edilmesini olanaklı kılan ilişkiler bulmaktır. Örnek olarak, “ev sahibi olan, evli, aynı iş yerinde beş yıldan fazladır çalışan, geçmiş kredilerinde geç ödemesi bir ayı geçmemiş bir erkeğin kredi skoru 825’dir.” Sonucu bir regresyon ilişkisidir.
- Öbekleme (Kümeleme): Kümeleme modellerinde amaç, küme üyelerinin birbirlerine çok benzediği, ancak özellikleri birbirlerinden çok farklı olan kümelerin bulunması ve veri tabanındaki kayıtların bu farklı kümelere bölünmesidir. Başlangıç aşamasında veri tabanındaki kayıtların hangi kümelere ayrılacağı veya kümelemenin hangi değişken özelliklerine göre yapılacağı bilinmemekte, konunun uzmanı olan bir kişi tarafından kümelerin neler olacağı tahmin edilmektedir. Örnek olarak bir süpermarketin müşteri bilgileri ve satış kayıtları incelenecek olursa, müşterilerin büyük bir kısmının düzenli olarak Cuma akşamları kredi kartıyla alışveriş yaptıkları şeklinde bir sonuca ulaşılabilir.
- Hipotez testi : Algoritmaları doğrulamaya dayalı algoritmalarıdır. Doğrulanacak hipotez VT üzerindeki verilerle belli doğruluk ve destek değerlerine göre sınanır. Sınama işlemi uzman tarafından aşağıdaki ihtiyaçlardan dolayı yapılır:
 - Bir kural ortaya çıkarılmak istendiğinde,
 - Ortaya çıkarılmış bir kuralın budanması veya genişletilmesinde.
- Varyans analizi : İkidenden çok kitle ortalaması arasındaki farkın önemini belirtir. Bir çok grubu aynı anda karşılaştırır.

ÖRNEĞİN; Üç yeni ilacın nabız atışı üzerindeki etkisi, çamaşır beyazlatması açısından değişik deterjanlar arasındaki fark varyans çözümlemesi ile incelenir. Eğer gruplar arasında farklılık çıkarsa ikişerli karşılaştırmalarla farkı yaratan grup belirlenebilir.

Çeşitleri:

- Tek yönlü varyans analizi
- Çift yönlü varyans analizi
- Tekrarlı ölçümlerde varyans analizi
- Çok etkenli varyans analizidir.

Varyans analizinin veri madenciliğinde kullanımı için, önceden veri tabanından belirli örnekleme yöntemleri kullanılarak seçilmiş olan verilerin analizinin yapılmasında kullanıldığı söylenebilir.

Varyans analizi, verinin yapısının belirlenmesinde ve ileriki aşamada model kurulmasında yol gösterici olarak kullanılacak bir istatistiksel yöntemdir.

- Lojistik regresyon : Doğrusal regresyonda Y (açıklanan) iki değer alan (yani; 0 ve 1) gösterge değişkeni olarak tanımlandığında bunlara ilişkin hata terimlerinin (e_i) beklenen değeri sıfır, $E(e_i) = 0$ ve varyanslarının sabit, $Var(e_i) = \sigma_e^2$ olduğu şeklinde tanımlanan varsayım gerçekleşmemektedir. Bunun bir sonucu olarak varsayımlardan sapma durumunda elde edilen tahminler en iyi doğrusal ve sapmasız tahmin ediciler olmayacaktır. Bu yetersizlik sınıflandırma analizlerinde doğrusal regresyonun kullanılmasını engellemektedir. Bu nedenle lojistik regresyon, sınıflandırma analizlerinde sık kullanılan yöntemlerden biridir. Lojistik regresyon, çok değişkenli normal dağılım varsayımına ihtiyaç göstermediğinden bu tür uygulamalarda üstünlük sağlamaktadır. Ayrıca sınıf üyeliğine ilişkin olasılıkları belirlemek özelliği de vardır. Lojistik regresyonun varsayımlarından biri doğrusal olasılık fonksiyonunun, hata terimlerinin dağılımının lojistik dağılıma uymasidir.
- X2 analizi: Ki-kare ilgi analizi pazarlama araştırmalarında çok yaygın olarak kullanılan bir istatistiksel analiz türüdür. Bu yaygın kullanımın en önemli nedenleri, çok basit bir analiz türü olması, varsayımlarının azlığı ve çok güçsüz ölçeklerde ölçülmüş verilere uygulanabilmesidir. İki veya daha fazla nitelik esas alınarak sınıflandırılan veriler değerlendirilerek bu nitelikler arasındaki ilginin derecesinin belirlenmesi (bağımsızlık testi) amacıyla kullanılır.

3.6.2. Bellek Tabanlı Teknikler

Bellek tabanlı veya örnek tabanlı bu yöntemler istatistikte 1950'li yıllarda önerilmiş olmasına rağmen o yıllarda gerektirdiği hesaplama ve bellek yüzünden kullanılamamış ama günümüzde bilgisayarların ucuzlaması ve kapasitelerinin artmasıyla, özellikle de çok işlemcili sistemlerin yaygınlaşmasıyla, kullanılabilir olmuştur. Bu yönteme en iyi örnek en yakın k komşu algoritmasıdır (k-nearest neighbor).

3.6.3. Genetik Algoritmalar

Diğer veri madenciliği algoritmalarını geliştirmek için kullanılan optimizasyon teknikleridir. Sonuç model veriye uygulanarak gizli kalmış kalıpları ortaya çıkarılmakta ve bu sayede tahminler yapılabilir. Doğrudan postalama, risk analizi ve perakende analizlerinde kullanılabilir.

3.6.4. Yapay Sinir Ağları

Bu yöntem, belirli bir profile uyuması için kalıp düzenlerini kontrol etmektedir ve bu süreç içerisinde belli bir öğrenme faaliyeti gerçekleştirerek sistemi geliştirmektedir. Yapay sinir ağlarında kullanılan öğrenme algoritmaları, veriden üniteler arasındaki bağlantı ağırlıklarını hesaplar. Yapay Sinir Ağları istatistiksel yöntemler gibi veri hakkında parametrik bir model varsaymaz yani uygulama alanı daha geniştir ve bellek tabanlı yöntemler kadar yüksek işlem ve bellek gerektirmez.

3.6.5. Karar Ağaçları

İstatistiksel yöntemlerde veya yapay sinir ağlarında veriden bir fonksiyon öğrenildikten sonra bu fonksiyonun insanlar tarafından anlaşılabilir bir kural olarak yorumlanması zordur. Karar ağaçları ise ağaç oluşturulduktan sonra, kökten yaprığa doğru inilerek kurallar yazılabilir. Bu şekilde kural çıkarma veri madenciliği çalışmasının sonucunun doğrulanmasını sağlar. Bu kurallar uygulama konusunda uzman bir karar vericiye gösterilerek sonucun anlamlı olup olmadığı denetlenebilir. Sonradan başka bir teknik kullanılacak bile olsa karar ağacı ile önce bir kısa çalışma yapmak, önemli değişkenler ve yaklaşık kurallar konusunda karar vericiye bilgi verir

IV. BÖLÜM : KURUMSAL PERFORMANS YÖNETİMİ

4.1. Kurumsal Performans Yönetimi

Performans; bir işi yapan bir bireyin, bir grubun ya da bir teşebbüsün o işle amaçlanan hedefe yönelik olarak nereye varabildiğinin, başka bir deyişle, neyi sağlayabildiğinin nicel (miktar) ve nitel (kalite) olarak anlatımıdır.

“Performans” en basit tanımıyla verimliliğin ölçülmesidir. Bu ölçme kurum için yapılırsa “Kurumsal Performans”, çalışanlara yönelik yapılırsa “personel performans değerlendirilmesi” amacı taşır ve işletmelerin personel politikasının etkinliğini ölçmede yarar sağlar.

İşletme performans değerlendirme sistemleri çalışanların belirli bir dönemdeki fiili başarı durumlarını ve geleceğe ilişkin gelişme potansiyellerini belirlemeye yönelik bir çalışmadır.

Performans Sisteminin Amacı:

Bir taraftan kurumun vizyonu doğrultusunda hedeflerini belirlemesi ve bu hedeflerin çalışanların katkılarıyla gerçekleşmesinin sağlanmasıyla, diğer taraftanda hedeflere ulaşırken çalışanların katılımlarının adil, sistemli ve ölçülebilir bir yöntemle değerlendirilmesi ve motive edici bir çalışma ortamı oluşturarak kişisel gelişimin desteklenmesidir.

4.2. Anahtar Performans Göstergeleri (KPI) Kavramı

Anahtar Performans Göstergeleri (KPI) ve “Anahtar Başarı Göstergeleri” olarak da bilinen yaklaşım, bir organizasyonun örgütsel hedeflerine yönelik ölçümlemesine ve tanımlamasına yardımcı olur.

Bir organizasyon, misyonunu analiz etmektedir, paydaşlarını tanımlamaktadır ve hedeflerini belirlemektedir; bu nedenle bu hedeflere yönelik ilerlemenin ölçülmesinde kullanılacak bir yola ihtiyaç duyulur ki bu ölçümler Anahtar Performans Göstergeleri’dir.

Anahtar Performans Göstergelerinin Özellikleri :

- Anahtar Performans Göstergeleri organizasyonun hedeflerine göre belirlenir, hedefler degistikçe performans göstergeleri de güncellenmelidir.
- Anahtar Performans Göstergeleri az sayıda olmalı ve departmanlarca benimsenmelidir.
- Anahtar Performans Göstergeleri ölçülebilir olmalıdır.
- Anahtar Performans Göstergeleri uzun dönemli olmalıdır.
- Anahtar Performans Göstergeleri tanımlanabilir olmalıdır.

Neden Ölçüm ?

- Ölçümleme başarının anahtarıdır.
- Odaklanma ve dikkat (ilgi) sağlar.

KPI-----→BI

Anahtar Performans Göstergeleri, faaliyetlerin öngörülmesi ve işletmenin şimdiki durumunun değerlendirilmesinde fonksiyonel departmanlar arasında stratejilerin yayılmasını sağlayan Kurumsal Zeka (Business Intelligence- BI) çalışmalarında kullanılır.

Stratejik Kurumsal Zeka ise kıdemli yöneticilere işletme ile ilgili gelişme fırsatları ve trendleri yakalamasını ve işletmeyi bütünsel bir bakış açısı ile görmesini sağlarken, aynı zamanda işletmede ortak bir çalışmayı teşvik etmek amaçlı departmanların sınırlarına karşı Anahtar Performans Göstergelerine yönelik şirketin denetlenmesinde de kullanılır.

Performans göstergeleri bir şirketin stratejisinin diğer şirketlerden farklı olan özelliklerini ortaya koyar. İyi bir Balanced Scorecard, hem sonuç ölçümleri hem de bu sonuçları elde etmeyi sağlayacak performans göstergelerini kapsamalıdır. İyi bir Scorecard, şirketin stratejisine özel olarak belirlenen sonuçlar ve performans göstergelerinin en uygun oranlardaki birleşimlerini kapsamalıdır. [Altun, 2001: 21]

4.3. Balanced Scorecard (Kurumsal Karne)

Balanced Scorecard, kurumların genel performansı ölçerken kullandıkları, ve aynı zamanda kişi ve departmanların günlük aktivitelerin kurumsal stratejiye paralel olup olmadığının takip edildiği bir yönetim metodolojisidir. Özellikle Kurumsal Karne projeleri olarak da nitelendirilebilen Balanced Scorecard yaklaşımının en önemli noktalarından birisi, bir kurumun genel performansını sadece finans perspektifindeki göstergelerle değil, aynı zamanda müşteri, iç süreçler, ve eleman perspektifinden inceleyebilmektir. Bu sırada farklı performans göstergeleri arasındaki ilişkilerin bir strateji haritası üzerinde birbiriyle nasıl etkileşim halinde olduğunun tasarlanması, bunların birbirine etkisinin ağırlıklandırılması ve gerektiğinde bir takım simülasyonların yapılabilmesi gerekir.

Balanced Scorecard projeleri geçmişte daha çok yönetim düzeyindeki küçük bir gruba hitap ederken, artık bir çok kurum tüm organizasyon bazına yaymaya çalışmaktadır. Bu durumda her çalışan sadece kendi karnesini görerek, kendi performansını da değerlendirebilmektedir.

Ağırlıklı olarak finansal muhasebe ölçümlerine dayanan performans ölçüm yöntemlerinin artık eskidigi ve geçerliliğini yitirdigi fikri yatmaktadır.

“Finansal”, “müşteri”, “şirket içi yöntemler (içsel)” ile “öğrenme ve büyüme” boyutları olan Kurumsal Karne, misyon ve stratejinin anlaşılması ve iletilmesi için bir çerçeve

ve bir lisan olusturur. Ölçüler yoluyla şirketteki tüm çalışanlara bugün ve gelecekte başarıyı elde etmeyi sağlayacak etkenler hakkında bilgi verir.

Karne'lerde yer alacak özel ölçüler seçmenin temel amacı, stratejiyi en iyi ifade edecek ölçüleri belirlemektir. Her strateji tek ve benzersiz olduğu için her Karne de tek ve benzersiz olmalı ve birkaç tane özel, kendine has ölçüye sahip olmalıdır.

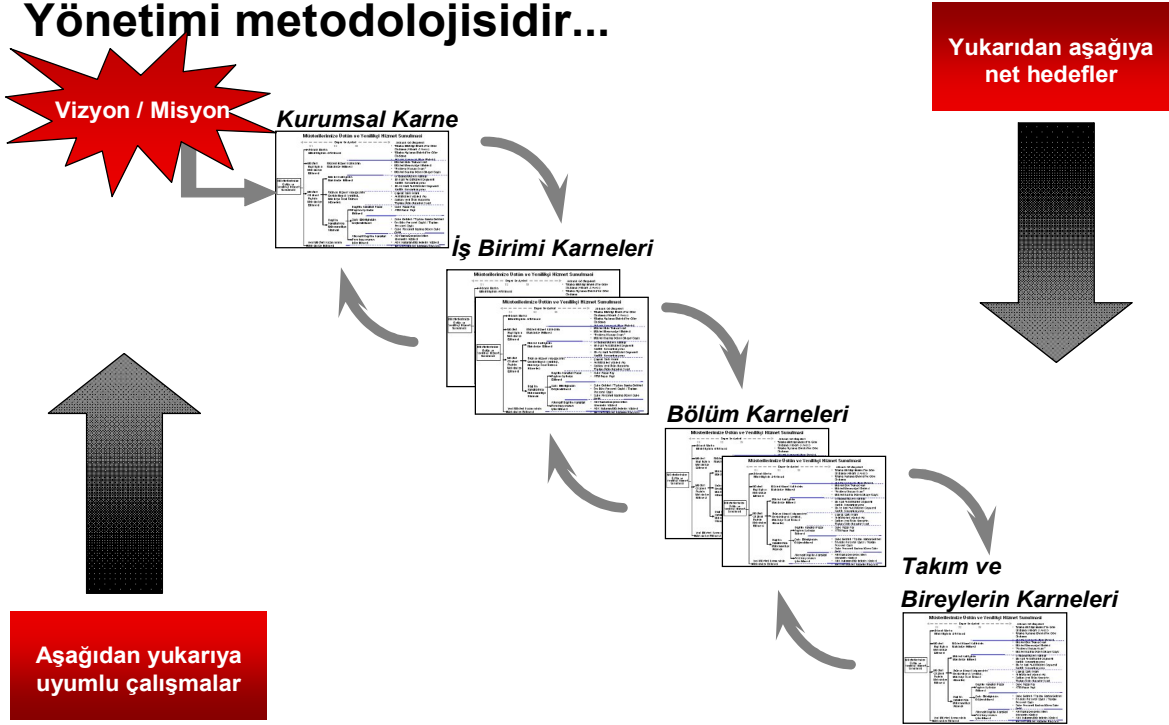
Balanced Scorecard, yöneticilere bugünkü faaliyetlerini yarınki hedefleriyle ilişkilendirmeleri imkanını verir.

Kurumlar Scorecard'ı şu amaçlarla kullanmaktadırlar:

- Stratejinin açıklığa kavuşturulması ve güncelleştirilmesi
- Stratejinin kurum geneline duyurulması
- Birimlerin ve bireylerin hedeflerinin strateji ile uyumlu hale getirilmesi
- Stratejik hedeflerin uzun dönemli hedefler ve yıllık bütçelerle bağlantısının kurulması
- Stratejik girişimlerin tanımlanması ve düzenlenmesi
- Strateji hakkında bilgi edinmek ve iyileştirmek amacıyla periyodik performans değerlendirmeleri yapılması.

Balanced Scorecard bir şirketin yönetim süreçlerini düzenlemesine yardımcı olur ve tüm kuruluşu uzun dönemli stratejinin uygulanmasına odaklar.

Balanced Scorecard; Kurumsal Performans Yönetimi metodolojisidir...



Şekil 4.1. Kurumsal Karne Basamakları

V. BÖLÜM : BANKACILIK SEKTÖRÜNDE ÖRNEK BİR İŞ ZEKASI UYGULAMASI

Bu bölümde bankacılık sektörü için hazırladığım Veri Ambarı, Veri Madenciliği, Analitik Modelleme ve Raporlamalar (Statik ve Analitik) üzerine tasarım, modelleme ve uygulama örnekleri açıklanmıştır. Yine bankacılık sektörü için hazırlanmış (tüm sektörler için uygulanabilir) İş Zekası Platform ve ETL tool seçim önerileri ve soru kümeleri verilmiştir.

İş Zekası projelerinin başarısı için dikkat edilmesi gereken noktalar ve proje yönetim önerileri anlatılmıştır. (Bankacılık Sektörü - Genele uygulanabilir)

5.1. Veri Ambarı Tasarım Özellikleri

Aşağıda banka sektörü için hazırladığım bir veri ambarı tasarım örneği verilmiştir. Bu örnekte geliştirilen veri ambarı özellikleri, datamart ve dimension içerik ve özellikleri ile uygulama alanları ve özellikleri aşağıda listelenmiştir.

1) Veri Ambarı (Data Warehouse) Özellikleri :

- Günlük Veri Ambarı (Veri ambarına veri akışının günlük olarak yapılması)
- Ortak Veri Ambarı (Tüm Uygulamalar için – Core Banking, Yatırım, Hazine, Kredi Kartları)
- Veri ile İlişkili Tüm Çalışma Sistemlerini Kapsama
- BI Katmanı ile verilen hizmetler;
 - Standart Raporlama
 - Analitik Raporlama ve MIS
 - Performans Analizi
 - Veri Madenciliği
- Veri Kontrolleri
 - Günlük ETL sürecinde
 - ❖ Kaynak ve Hedef Kayıt Sayısı Eşitliği Kontrolü
 - ❖ Ana ve Boyutsal Tabloların Foreign Key kullanılarak tutarlılıklarının kontrolü
 - ❖ Defter-i Kebir Bilanço / Hesap Bilanço
 - ❖ Hesap Balans/Günlük İşlem Miktarı
 - ❖ MIS Şubesi (Gözetim ve Resmi Raporlamalar)
 - ❖ Ürün Tanımı (Ürün Hiyerarşisi)
 - Finansal Kontrol
 - ❖ Varlık – Taahhüt Farklılaşması
 - ❖ Mevduat Hesapları üzerindeki Hesap Açığı Bilançosu
 - ❖ Belirli Defter-i Kebirler arasındaki eşitlik prensipi denetimi
 - ❖ İptal Edilmiş Muhasebe Kayıtları

2) Datamart ve Dimensionların Özellikleri:

Müşteri Bilgileri

- Banka ve tüm iştiraklerdeki (yurtdışı temsilcilikler dahil) müşterilerin unique müşteri numarası ile tutulması
- Müşterinin ait olduğu sektör bilgisi ve bağlı alt sektör bilgisi
- Sırası ile ;
 - Portföy (Portföy Yöneticisi bilgisi dahil),
 - İşkolu bağlantılarının kurulabilmesi (Segmentel)
- Demografik bilgiler,
 - Yasal statü,
 - Kişisel bilgiler
 - Adres bilgileri
 - İletişim bilgileri
 - Yasal bildirim amaçlı bilgiler
- Müşteri davranışı ile ilgili segmentasyon sonuçları ve tanımlamalar
 - Kampanyalar
 - Yapılabilecek kampanyalar,
 - Yapılan kampanyalar,
 - Kampanya sonuçları
 - Ürün eğilimleri
 - Ürün sahipliği
 - Ürün eğilimleri
 - Eğilim sonrası ürün alma durumu
 - Churn etme ihtimalleri
 - Ürün bazlı rating
 - Müşteri bazlı rating
 - Basel 2 rating

Ürün Yapısı

- Banka ve tüm iştirakler detayında (yurtdışı temsilcilikler dahil) ürün eşleştirilmesi yapılmış olarak, (Tek bir ürün yapısı)
- Temel ürün yapısı
 - En temelde müşterinin de altında kontrata bağlanabilen,
- İhtiyaca göre temel ürünlerden ürün grupları oluşturma
 - Yasal raporlamalar ,
 - MIS (Management Information System) raporlamalar ,
 - Operasyonel raporlamalar,
 - Karar destek sistemleri

Limitler

- Banka ve tüm iştirakler detayında (yurtdışı temsilcilikler dahil),
- Döviz kısıtlımları olan, (Döviz cinsi bazında)
- Sırası ile
 - Müşteri grubu
 - Müşteri
 - Ürün grubu
 - Ürün kısıtlımlarında
- History
 - Değişim/Revizyonlar,
 - Takibe alınma tarihi itibari ile
 - Sorunlu hale dönüş itibari ile
- Kredi ve kontrat ilişkisi ile,
- Kullanım şartlarını ihtiva edecek şekilde,
- Limit boşlukları hesaplanacak şekilde risk ilişkisi ile

Riskler

- Banka ve tüm iştirakler (yurtdışı temsilcilikler dahil) unique müşteri numarası ile (Tüm Grup bazında Risk tanımı)
- Özel risk grupları bazında, farklılaşan detay özellikleri barındıran;
 - Nakit riskler
 - ◆ Kredi kartları (kart bazında)
 - ◆ İskonto kredileri
 - ◆ Taksitli krediler
 - ◆ Taksitli DEK
 - ◆ DEK (Döviz Endeksli Krediler)
 - ◆ Kıymetli maden kredileri
 - ◆ Eximbank kredileri
 - ◆ Nakit karşılıklı krediler
 - Gayri nakit riskler
 - ◆ Teminat mektupları
 - ◆ Akreditifler
 - ◆ Taahhütler
- Sırası ile
 - Konsolide (iştirakler dahil)
 - Banka Solo
 - Müşteri grubu
 - Müşteriler
 - Temel ürünler bazında

- Kontrat bazında kırılımda
- Teminat ilişkisi kurulmuş, sırası ile
 - Konsolide (iştirakler dahil)
 - Müşteri grubu
 - Müşteri,
 - Ürün
 - Kontrat bazında
- Limit ilişkisi kurulmuş, sırası ile
 - Konsolide (iştirakler dahil)
 - Müşteri grubu
 - Müşteri,
 - Ürün
 - Kontrat bazında

Kredi Kartları

- Unique müşteri numarası detayında,
- Kart bazında
- Kart türü bazında
- Statüler,
- Süreçler,
 - Talep,
 - Talep/satış kanalı,
 - Basım ,
 - Teslim,
 - Yenileme,
 - Limit,
- Performans,
- Kampanya,
- Karlılık,

Teminatlar

- Banka ve tüm iştirakler (yurtdışı temsilcilikler dahil) unique müşteri numarası ile (Grup Bazında)
- Bankanın teminat yapısı ile uyumlu,
- Basel kriterleri ile uyumlu detayda (CRM= Credit Risk Mitigation),
- Karşılıklar kararnamesi ile uyumlu detayda,
- Sırası ile
 - Konsolide (iştirakler dahil)
 - Banka Solo

- Müşteri grubu
- Müşteriler
- Temel ürünler bazında
- Kontrat bazında kırılımda
- Risk ilişkisi kurulmuş, sırası ile
 - Konsolide (iştirakler dahil)
 - Müşteri grubu
 - Müşteri,
 - Ürün
 - Kontrat bazında
- Limit ilişkisi kurulmuş, sırası ile
 - Konsolide (iştirakler dahil)
 - Müşteri grubu
 - Müşteri,
 - Ürün
 - Kontrat bazında

Çekler

- Bankamız çekleri,
- Diğer banka çekleri,
- Karşılıksız çekler,
- Tahsil çekleri
- Teminat çekleri (teminat modül ve bağlantılı risk ilişkisi kurulmuş olarak)
- Yasal çek bildirimleri
 - TCMB
 - Gözetim seti

Senetler

- Tahsile alınan
- Teminata alınan (teminat modül ve bağlantılı risk ilişkisi kurulmuş olarak)

Muhasebe Hareketleri

- Transaction bazında (Bağlı grup ilişkisi ile birlikte),
- Hesap detayında,
- Kanal bilgisi ile
- Karşı hesap bilgisi ile,
- Bağlı Defteri Kebir ve müşteri hesabı ilişkisi ile,

- History,

3) Uygulama ve Fonksiyonların Özellikleri :

Operasyonel ve Anlık (Adhoc) Raporlar

- Fonksiyonlar
 - 7/24 kesintisiz erişim,
 - Standart raporlarda özelleştirme fonksiyonu ile rapor çeşitlendirebilme
 - Özelleştirilen raporun kullanıcı bazında muhafazası ve yeniden çalıştırılabilmesi (Template oluşturma ve çalıştırabilme özelliği)
- Period
 - Günlük
 - Haftalık,
 - Aylık
 - Çeyrek dönemlik
 - Yıllık,
 - Set edilen belirli tarihler bazında(örneğin: her ayın 25.günü , her haftanın 2. günü vb..)
- Microsoft entegrasyonu / Değişik formatlardaki tablolara bilginin alınması
 - Excel
 - Text
 - Powerpoint
 - Vb.
- Kullanıcı rolüne göre uygulama erişimini sınırlama (Security)

Ajanda Uygulaması

- Farklı kullanıcı grupları için;
 - 7/24 kesintisiz erişim,
 - Günlük izlenmesi gereken detayları / uyarıları / takipleri
 - ❖ Ürün kırımında tüm ürünler ve müşteriye ait bilgiler için,
 - ❖ Vade,
 - ❖ Taksit,
 - ❖ Doğumgünü
 - Kullanıcı rolü bazında
 - Detay ve özet olarak izleyebilmek
- Period
 - Günlük
 - Haftalık,
 - Aylık

- Çeyrek dönemlik
- Yıllık,
- Set edilen belirli tarihler bazında(örneğin: her ayın 25.günü , her haftanın 2. günü vb..)
- Microsoft entegrasyonu / Değişik formatlardaki tablolara bilginin alınması
 - Excel
 - Text
 - Powerpoint
 - Vb.
- Kullanıcı rolüne göre uygulama erişimini sınırlama

Bütçe ve Hedef Uygulamaları

- Detayı - Sırası ile ;
 - Müşteri,
 - Müşteri grupları,
 - Müşterilerden oluşan portföyler,
 - Portföylerden oluşan şube-işkolları (Segment),
 - Şubeler,
 - Şubelerden oluşan bölgeler ve bölgedeki işkolları,
 - Banka bazında işkolları,
 - Banka konsolide hedef ve bütçeler
- Hacimsel hedefler,
- Adetsel hedefler,
- Getiri hedefleri,
- Masraf bütçesi,
- Fonksiyonlar;
 - Giriş,
 - Güncelleme,
 - Onaylama,
 - Upload,
 - İzleme,
 - Kontrol ve otomatik hesaplamalar,

MIS ve Karar Destek Sistemleri

- Banka ve tüm iştirakler (yurtdışı temsilcilikler dahil) unique müşteri numarası ile
- 7/24 erişim ve izleme,
- 2 tarih arası karşılaştırma,
- Zaman serilerinde gelişim izleme,

- Detayı Sırası ile ;
 - Kontrat bazında,
 - Ürün bazında,
 - Müşteri bazında,
 - Müşteri grupları bazında,
 - Portföy bazında,
 - Şube-işkolu bazında,
 - Şube bazında,
 - Bölge bazında,
 - İşkolu bazında,
 - Banka konsolidesinde / Kampanya bazında,
 - Farklı uygulamalar için temel ürünlerden yola çıkarak oluşturulmuş farklı ürün grupları bazında,
 - Kanal bazında,
 - Alt detay yada hesaplanmış / gruplanmış özetler bazında,
 - ❖ İşlem sayıları,
 - ❖ Risk,
 - ❖ Ortalama,
 - ❖ Nokta bakiye,
 - ❖ Tüm döviz cinsleri karşılığında,
 - ❖ Durasyon,
 - Segment bazında,
 - Sektör bazında,
- Period
 - Günlük,
 - Haftalık,
 - Aylık,
 - Çeyrek dönemlik,
 - Yıllık,
 - Yıllık Kümülatif
- Detay bazda İlişkisel karşılaştırma,
 - Fiili,
 - Hedef/bütçe,
 - Kar/zarar,
 - ❖ Getiri,
 - ❖ Maliyet
 - Gerçekleşme,
 - Transaction Cost
 - Funding Cost – Transfer Pricing

Verilen tüm bu özellikler çalışmamız kapsamında ortaya çıkarılmıştır ve düzenlenmiştir.

5.2. Veri Madenciliği Uygulamaları

Aşağıda çalışmalarımız sonucunda elde ettiğimiz veri madenciliğinin önemli noktaları ve banka sektöründe kullanılan veri madenciliği modellemeleri, amaçları ve özellikleri listelenmiştir.

Önemli Noktalar:

- ✓ Datamartların oluşturulması (Amaca Yönelik)
 - Değişkenlerin belirlenmesi
 - Hedefe yönelik değişkenlerin belirlenmesi
 - Belirlenen değişkenlerin data kalitesinin ve doluluk oranlarının tespiti.
- ✓ Öğrenme Kümesinin belirlenmesi
 - Periyodu
 - Müşteri sayısı
 - Öğrenme kümesindeki hedef müşteri gruplarının adetsel dağılımı (Balans)
 - Kümeleme Yapılıp Yapılmayacağı (Belli kümeler/müşteri grupları için tek model geliştirilebilir)
 - Modelin öğrenme periyodu,
 - Modelin test periyodu
- ✓ Modellemede kullanılacak değişkenlerin belirlenmesi
 - Öğrenmede etkili olacak değişkenlerin belirlenmesi
 - Boş ve outlier alanların ne değer alacağını belirlenmesi, nümerik alanların gerekirse set alanlara dönüştürülmesi
- ✓ Modellemede kullanılacak algoritmanın belirlenmesi, expert ayarlarının yapılması
 - Regresyon (doğrusal, doğrusal olmayan, lojistik)
 - Kümeleme Algoritmaları
 - Karar Ağaçları
 - Yapay Sinir Ağları
 - Genetik Algoritmalar
 - Bellek Tabanlı Algoritmalar
- ✓ Oluşturulan modellerin performanslarının ölçülmesi/kıyaslanması

Performans Kriterleri:

- İsabet oranını (hit rate)
- Kapsama oranı (capture rate)
- Kaldıraç etkisi (lift)

- Ortalama skorun gerçek batık oranına yakınlığı
 - Öğrenme yüzdeleri (accuracy)
- ✓ Oluşturulan modellerin performanslarının kıyaslanması sonucunda kullanılacak modele karar verilmesi.

5.2.1. Ürün Eğilim Modellemeleri

Farklı yöntemler (veri madenciliği, kural bazlı) ve metodlar ile müşterilerin almaya eğilimli oldukları ürünlerin belirlenmesi ürün eğilim modellemeleri ile gerçekleştirilir.

- Müşterilerin bağlı oldukları pazarlama gruplarına göre farklı modellerin oluşturulması
- Ürün alma eğilimi olan müşterilerin kampanya yönetimi uygulamasına çıkılarak kampanyaların başlatılması
- Müşteri CRM ekranlarında müşterilerin eğilimli oldukları ürünlerin çıkılarak izleme/sorgulama yapılabilmesi
- PY lerin belli bir ürünü almaya eğilimli olan müşterilerini sorgulayabilmesi, gelen liste üzerinden müşteriye arayabilmeleri
- Gişe Satış modülüne müşterilerin eğilimli oldukları ürünlerin çıkılması
- Belirli tarihlerde modellerin tekrar çalıştırılarak, güncel eğilimlerin sisteme çıkılması
- Oluşturulan listeler için kullanıldıkları uygulama bazında performans takibinin yapılması

5.2.1.1. Çapraz Satış (Cross Sell)

Çapraz Satış, Müşterilerin demografik bilgilerinin, ürün sahipliklerinin ve kullanımlarının veri madenciliği yöntemleri ve istatistikî metodlar ile analiz edilerek sahip olmadıkları ancak almaya eğilimli oldukları ürünlerin belirlenerek satış faaliyetlerinde bulunulması yöntemine denir.

- Modellemelerde kaynak olarak kullanılacak pazarlama data martlarında yer alacak alanların/değişkenlerin belirlenmesi
- Data martların aylık olarak oluşturulması
- Herbir ürün için veri madenciliği yöntemleri kullanılarak müşterilerin ilgili ürünü alma olasılıklarının belirlenmesi (kredi kartı, TL vadeli hesap, YP Vadeli hesap, fon, internet şube, bireysel kredi, çek, otomatik ödeme talimatı, YP vadesiz hesap)
- Data martlar üzerinden kural bazlı da olarak ürün almaya eğilimli müşterilerin belirlenebilmesi

5.2.1.2. Ek Satış (Up Sell)

Müşterilerin demografik bilgilerinin, ürün sahipliklerinin ve kullanımlarının veri madenciliği yöntemleri ve istatistiki metodlar ile analiz edilerek, sahip oldukları ürünler içinde kullanımı artırmaya eğilimli oldukları ürünlerin belirlenerek ek satış faaliyetlerinde bulunulması yöntemidir.

- Aylık olarak oluşturulan datamartlar üzerinden veri madenciliği yöntemleri kullanılarak, müşterilerin ek satış eğilimleri belirlenir.
- Data martlar üzerinden kural bazlı da olarak ek kullanıma eğilimli olan müşterilerin belirlenebilmesi
- Müşterilerin bağlı oldukları pazarlama gruplarına göre farklı modellerin oluşturulması

5.2.1.3. Dizi Analizi (Sequence analysis)

Müşterilerin ürün alma sıralarının veri madenciliği yöntemleri ve istatistiki metodlar ile müşteri ürün alım sıralarında kalıpları analiz edilerek müşterinin bir sonraki alacağı en olası ürünün belirlenerek satış faaliyetlerinde bulunulması yöntemidir.

- Müşterilerin ürün alma sıralamaları ve hangi tarihlerde aldıkları bilgilerinin yer aldığı dizi data martlarının oluşturulması
- Oluşturulan datamartların aylık olarak beslenmesi
- Data martlar üzerinden kural bazlı da olarak ek kullanıma eğilimli olan müşterilerin belirlenebilmesi

5.2.1.4. Sepet Analizi (Market-Basket analysis)

Müşterilerin sahip oldukları ürünlerin veri madenciliği yöntemleri ve istatistiki metodlar ile analiz edilerek, ürün birlikteliklerinin tespit edilmesi yöntemidir.

5.2.2. Kümeleme

Birbirine benzer verileri ortak gruplarda toplayarak diğerlerinden ayırma yöntemidir.

- Veri madenciliği yöntemleri kullanılarak data martlar üzerinden kümeleme çalışmalarının yapılması

5.2.3. Terk Analizi (Churn)

Bankayı terk edecek müşterilerin önceden tespit edilmesi ve veri madenciliği yöntemleri kullanılarak müşterilerin terk eğilimlerinin belirlenmesidir.

- Terk analizlerinde kaynak oluşturacak churn data martlarının oluşturulması
- Datamartların aylık olarak update edilmesi
- Terk eğilimli müşterilerin terk etme olasılıklarına göre gruplara ayrılması

- Müşteri CRM ekranlarında müşterilerin terk etme olasılığı seviyelerinin görüntülenmesi
- PY lerin terk etme eğilimindeki müşterilerini terk seviyelerine göre sorgulayabilmesi
- Terk etme eğilimindeki müşterilerin listesinin kampanya yönetimi uygulamasına entegre edilmesi
- Belirli tarihlerde modellerin tekrar çalıştırılarak, sisteme çıkılması
- Oluşturulan listeler için performans takibinin yapılması

5.2.4. Davranışsal Skorlama (Behavioral scoring)

Müşterilerin davranışlarının (ürün sahipliklerinin, kullanımlarının ve ödeme performanslarının) veri madenciliği yöntemleri ve istatistiki metodlar ile analiz edilerek, takip eden bir yıl içinde batma (yasal takibe düşme) olasılıklarının belirlenmesi yöntemidir.

Müşteriler için üretilen davranışsal skor bilgisi aşağıdaki amaçlar için kullanılabilir.

- Davranışsal sorlama analizlerinde kaynak oluşturacak davranışsal skorlama data martlarının oluşturulması
- Datamartların aylık olarak update edilmesi
- Veri madenciliği yöntemleri ile bireysel ve SME müşterilerin batma olasılıklarının tespit edilmesi
- Veri madenciliği yöntemleri ile ürün bazlı batma olasılıklarının belirlenmesi (kredi kartı, bireysel kredi, KMH)
- Müşterilerin batma olasılıklarına göre risk seviyelerinin belirlenmesi
- Batma olasılığı ve risk seviyelerinin diğer uygulamalara (müşteri veri tabanı, kredi ürünü başvuru süreci..) çıkılması
- Kredi ürünleri ile ilgili operasyonel ve uyarı raporlarına batma olasılığı ve seviyesi bilgilerinin eklenmesi,
- Müşterilerin batma olasılıklarının kredi başvuru skorkartında input olarak kullanılması
- Müşterilerin batma olasılıklarının limit arttırımı/azaltımı/ otomatik vade uzatımı kararlarında input olarak kullanılması
- Müşterilerin batma olasılıklarının karşılık ayırma işleminde input olarak kullanılması (Basel II)
- Ödemesini geciktiren müşterilerin aranmasında farklı aksiyonların alınması
- Müşterilerin davranışsal skorlarının risk bazlı fiyatlamada input olarak kullanılması

5.2.5. EAD hesaplamaları

Müşterilerin kredi ürün sahiplikleri ve davranışlarının veri madenciliği yöntemleri ve istatistiki metodlar ile analiz edilerek, EAD (Exposure at default - batık tutarı) değerinin tespit edilmesidir.

- EAD analizlerinde kaynak oluşturacak data martların oluşturulması
- Datamartların aylık olarak update edilmesi
- Veri madenciliği yöntemleri ile bireysel ve SME müşterilerin EAD değerlerinin tespit edilmesi
- Hesaplanan EAD değerlerinin kredi raporlarına eklenmesi

5.2.6. LGD hesaplamaları

Müşterilerin kredi ürün sahiplikleri ve davranışlarının veri madenciliği yöntemleri ve istatistiki metodlar ile analiz edilerek, LGD (Loss Given Default – batıktan kurtarılan tutar) değerinin tespit edilmesidir.

- LGD analizlerinde kaynak oluşturacak data martların oluşturulması
- Datamartların aylık olarak update edilmesi
- Veri madenciliği yöntemleri ile bireysel ve SME müşterilerin LGD değerlerinin tespit edilmesi
- Hesaplanan LGD değerlerinin kredi raporlarına eklenmesi

5.2.7. CCF hesaplamaları

Müşterilerin kredi ürün sahiplikleri ve davranışlarının veri madenciliği yöntemleri ve istatistiki metodlar ile analiz edilerek, CCF (Credit Conversion Factor – Risk ağırlığı) değerinin tespit edilmesi yöntemidir.

- CCF analizlerinde kaynak oluşturacak data martların oluşturulması
- Datamartların aylık olarak update edilmesi
- Veri madenciliği yöntemleri ile bireysel ve SME müşterilerin risk ağırlıklarının hesaplanmasında kullanılacak CCF değerlerinin tespit edilmesi
- Hesaplanan CCF değerlerinin kredi raporlarına eklenmesi

5.3. Analitik Modelleme ve Raporlama Uygulamaları

Aşağıda banka sektörü için hazırlanmış analitik modeller, tasarımları, özellikleri ve oluşturulan raporlama/izleme (Gişe ve portföy yönetimine yansıma) örnekleri gösterilmiştir.

5.3.1. Segmentasyon

Müşterilerin demografik bilgileri, ürün sahiplikleri, ürün kullanımları, varlık ve kredi bilgilerine göre gruplandırılması ile modellenir.

- Segmentasyona kaynak olacak verilerin (ürün sahiplik, hacim, kullanım, aktiflik...) aylık/3 aylık olarak sistemde data martlarda tutulması
- Periyodik olarak, 3 aylık data martlardaki veriler üzerinden; müşterilerin ait oldukları pazarlama gruplarına göre kural bazlı segmentlerinin belirlenmesi
- Belirlenen segment bilgilerinin müşteri veri tabanına ve diğer raporlamalara çıkılması
- Müşteri CRM ekranlarında müşterilerin segment bilgisinin izlenmesi
- Portföy Yöneticisi, Bölge müdürlükleri, Genel Müdürlük dashboardlarında segment kırılımında müşterilerin izlenmesi
- Segment bazında ürün sahiplik, kullanım ve hacim bazında izlemelerin yapılabildiği segment karnelerinin periyodik oluşturulması

Segment bazında:

- Müşteri Adedi,
- Aktif Müşteri Adedi
- Çalışma Büyüklüğü (ortalama)
- Toplam Kaynak + Yatırım (ortalama)
- Toplam Kredi (ortalama)
- Toplam Ürün adedi
- Kaynak + Yatırım ürünleri için ürün bazında; (Vadeli, Vadesiz, Fon A....)
 - Segmentteki ürünü aktif kullanan müşteri adedi,
 - İşkolundaki ürünü aktif kullanan müşteri adedi
 - Segmentteki ürünü aktif kullanan müşterilerin toplam hacmi
 - İşkolundaki aktif müşteri adedi içindeki oranı (%)
 - İşkolundaki aktif müşterilerin toplam hacmi içindeki oran (%)
- Kredi ürünleri için ürün bazında; (KMH, Bireysel Kredi, Ticari Kredi,...)
 - Segmentteki ürünü aktif kullanan müşteri adedi,
 - İşkolundaki ürünü aktif kullanan müşteri adedi
 - Segmentteki ürünü aktif kullanan müşterilerin toplam hacmi
 - İşkolundaki aktif müşteri adedi içindeki oranı (%)
 - İşkolundaki aktif müşterilerin toplam hacmi içindeki oran (%)
- Adetsel ürünler için ürün bazında; (KK, POS, Çek, internet şube...)
 - Segmentteki ürünü aktif kullanan müşteri adedi,
 - İşkolundaki ürünü aktif kullanan müşteri adedi
 - Segmentteki ürünü aktif kullanan müşterilerin toplam hacmi

- İşkolundaki aktif müşteri adedi içindeki oranı (%)
- İşkolundaki aktif müşterilerin toplam hacmi içindeki oran (%)
- Müşterilerin Segment bilgilerine göre PY ler arasında müşteri transferlerinin otomatik gerçekleştirilmesi
- Düşük segmentteki müşterilerin otomatik olarak havuza atanması

5.3.2. Hesap İşletim Ücreti

Müşterilerden periyodik olarak alınacak hesap işletim ücretlerinin belirli kurallara göre belirlenerek tahsil edilmesi işlemidir.

- Pazarlama grubu bazında kriterlerin ve tahsil edilecek tutarın farklılaşması
- Datawarehouse da müşterilerin muafiyet tanımları, ürün kullanımları, varlıkları, kredileri ve karlılık bilgilerine göre kural bazlı tahsilat listelerinin oluşturulması
- Üye işyerleri için pos tipi ve ciro bilgilerine göre tahsilat listelerinin oluşturulması, Pos tipine göre tahsil edilecek tutarın farklılaştırılması
- Datawarehouse da periyodik olarak oluşturulan tahsilat listelerinin oluşturularak tahsilat yapılmak üzere temel bankacılık uygulamasına aktarılması
- Müşteri muafiyet tanımının dikkate alınarak alınacak tutarın farklılaşmasının sağlanması
- Tahsilat yapılamaması durumunda belirli bir süre daha sistem tarafından kontrol yapılarak tahsilat işleminin tekrar denemesi
- Kısmi tahsilat yapılması
- Hesap işletim tahsilatlarının raporlanması
- Muaf müşterilerin raporlanması
- Hesap işletim tahsilatlarının müşteri karlılığı hesaplamalarında kullanılması

5.3.3. Gişe Satış Modülü

Müşterilerin eğilimli olduğu ürünler için gişe ekranlarında uyarı mesajlarının görüntülenmesi işlemidir.

- Veri madenciliği yöntemleri veya kural bazlı belirlenen ürün alma eğilimli müşteriler listesinin gişe satış modülüne çıkılması
- Gişe işlemi (nakit yatırma, nakit çekme, vb...) sırasında kullanıcı işlemi tamamlamak istediğinde; ekranda müşterinin almaya eğilimli olduğu ürünlerin görüntülenmesi
- Kullanıcı listeden herhangi bir ürünü seçtiğinde; ürün için tanımlanmış olan öneri cümlesinin ekranda görüntülenmesi

- Müşterinin ilgilenmiyorum demesi durumunda bu bilginin sisteme alınarak, sonrasında müşteri için ilgili ürün eğiliminin görüntülenmesinin engellenmesi
- Müşterinin ilgilenmesi durumunda; kaydın portföy yöneticilerinin ulaşabildiği gişe müşterisi eğilim havuzuna yönlendirilmesi
- Müşterisinin istemesi üzerine PY ye yönlendirilen ürünler için; sonrasında müşteriye ilgili ürünün eğiliminin görüntülenmesinin engellenmesi
- Portföy yöneticilerinin, havuzdaki kayıtları izlemelerinin sağlanması
- Portföy yöneticilerinin havuzdan kayıt seçerek, ürün satışı ile ilgili Başarılı, Başarısız aksiyonlarının girilmesinin sağlanması
- Portföy yöneticilerinin havuzdaki herhangi bir kaydı Beklemede statüsüne almalarının sağlanması
- PY havuzunda bekleyen kayıtların belirli bir gün geçtikten sonra direkt izlenmesinin engellenmesi
- Gişe Satış Modülünde görüntülenen ürün alma eğilimlerinin performans takiplerinin ve raporlamasının yapılması
- Gişe Satış modülünden yönlendirilen ve satışı gerçekleştirilen ürünler için; gişeciye belirli bir oranda prim verilmesinin sağlanması
- Belirli periyotlarda gişe satış modülünde yer alan ürün eğilim listelerinin görüntülenmesi

5.3.4. Müşteri Sadakat Uygulaması (Loyalty)

Aylık olarak müşterilerin ürün kullanımları ve yeni sahipliklerine dayanan ve aşağıdaki sistemleri kapsayan ödül uygulamasıdır.

- Ürün kullanımları üzerinden puan kazanım sistemi
- Bankacılık sistemi ile kredi kartları tarafında kazanılan puanları kullanım sistemi
- Puan kazanım ve kullanımının izlenmesine yönelik raporlama sistemi

Müşterilerin puan kullanım tercihlerinin/değişikliklerinin tüm kanallardan alınması

Veri Ambarı üzerinde aylık puan hesaplanması

- İşkolu bazında puan hesaplanacak müşteri segmentlerinin belirlenmesi
- Puan hesaplanacak ürünlerin işkolu bazında belirlenmesi
- Sistemde işkolu ve ürün bazında aşağıdaki puan hesaplama yöntemlerinin parametrik tutulması
 - İlk kullanımda puan verilir verilmeyeceği bilgisi
 - İlk kullanım için puan verilecekse; verilecek puan

- Ay içindeki kullanımda sabit puan mı, katsayı mı verileceği bilgisi
 - Puan hesaplanırken, kullanım tutarı üzerinden katsayı ile çarpım yapılacaksa kullanılacak katsayı
 - Puan hesaplanırken sabit kullanım puanı verilecekse, verilecek sabit puan
- Min. Kullanım tutarı (puan hesaplanması için gereken min. Kullandırım tutarı)
- Ürün için müşteriye verilecek Max. puan değeri
- Ürün bazında hesaplanan puanların toplanarak müşteri toplam puanının hesaplanması
- Müşteri toplam puanının kullanılacağı yere göre (para (fon, masraf/komisyon indirimi) ya da bonus puan) kullanılabilir tutarın/puanın belirlenmesinde; işkolu bazında tanımlanan katsayıların kullanılması

Puan Kullandırım Sistemi :

- Kazanılan puanların işkolu bazında farklı uygulamalarda kullanılabilmesi
- Bonus Puan
- Müşterinin yaptığı bazı bankacılık işlemlerinde masraf alınmaması (kazanılan puan kadar)
- Fon Alışı
 - Alınan fonun kullanımı ile ilgili kısıtlama olmaması
 - Alınan fonun sadece Kredi Kartı ödemesinde kullanılması
- Belirli bir süre kullanılmayan puanların silinmesi

Puan Raporlama:

- Aylık kazanılan puanların raporlanması
- Ay içinde kullanılan puanların raporlanması

5.4. Analitik ve Statik Raporlar

Aşağıda banka sektörü için iş zekası üzerinden hazırlanan raporlama çeşitleri gösterilmiştir.

Verim Yönetimi (Hacim, Verim ve Karlılık Analizi):

- Müşteri/Portföy Yöneticisi/Şube/Bölge/İş kolu ve Banka bazında Gelir/Gider Rasyoları, Günlük/Aylık/Yıllık ve Üçer aylık dönemler itibariyle kümüle hacim, getiri takibi ve detay bazlı Gelir/Gider Tablosu,
- Ürün bazlı kar analizi,
- Faiz dışı gelirler,
- Operasyonel maliyet analiz tabloları.

Finansal İş Takibi:

- Günlük geri dönüşler
- Taksit ödemeleri
- Fatura tahsilatları

MIS:

- Ürün bazında İş birimi/şube /bölge/banka kırılımında hacim ve adetsel gelişim raporları
- Vadeli Mevduat Analizi
- Vadesiz mevduat ve Maliyetsiz Kaynak Analizi
- Çek tahsilatları ve Takas Hacimleri
- Nakdi/G.Nakdi Kredi Analizi
- Para Transferleri(EFT/Havale) İşlem hacim raporları
- Şube/Kullanıcı bazlı İşlem adetleri
- Ürün bazlı kurumsal performans raporları(Tüm ürünlerin hacim ve adetsel gelişiminin çeşitli boyutlarda izlenebildiği özet analizler)
- Şube ve satış temsilcisi bazında hedef gerçekleştirme raporu

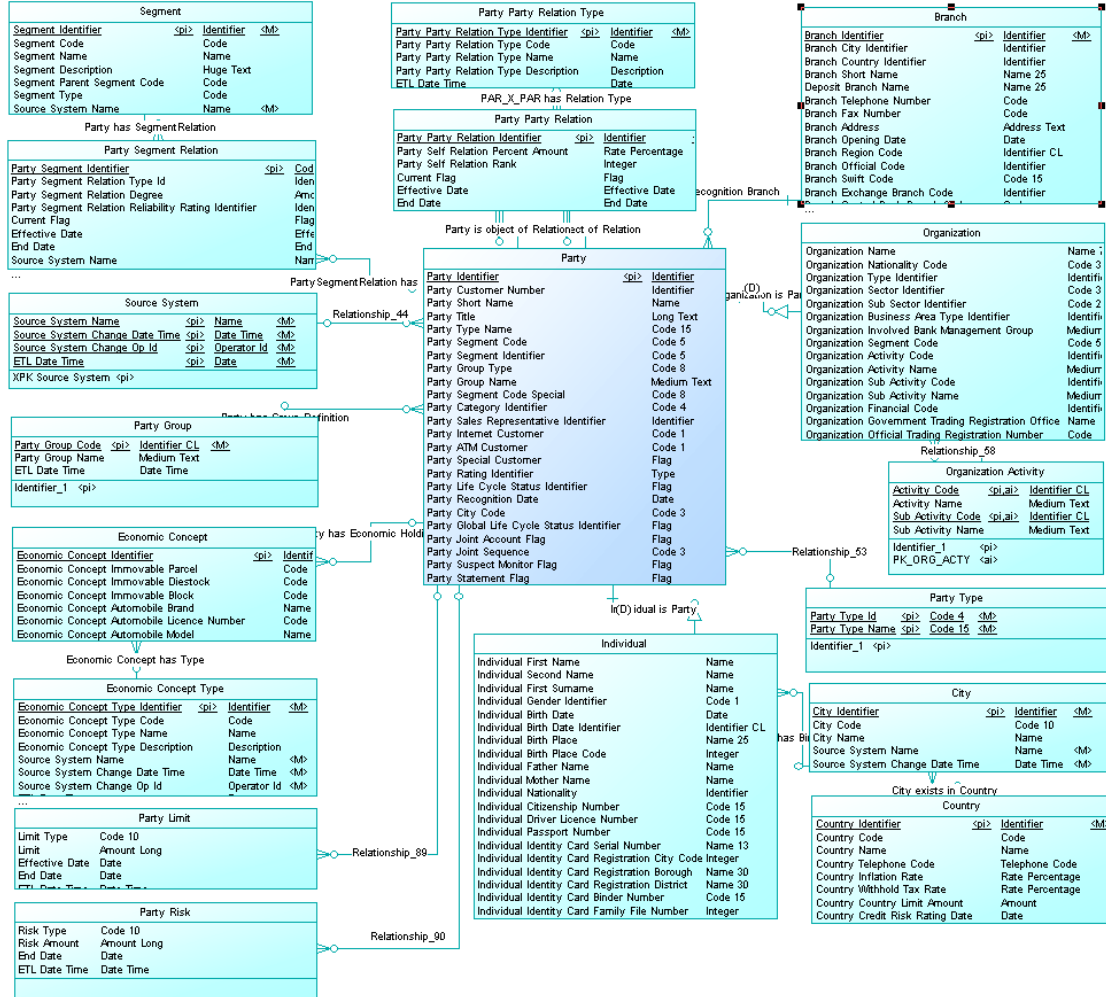
Kurumsal Performans Yönetimi:

- Kurumsal Bütçe yönetimi, işkolları, şube ve portföy yöneticilerine verilmiş hedefler ve gerçekleştirme analizleri.
- Satış faaliyetlerinden elde edilen hacim ve gelirin portföy yöneticisi bazında ölçümlenebilir hale getirilerek bütçe hedeflerinin tüm kademelerde gerçekleştirme analizleri.

5.5. Veri Ambarı Modelleme

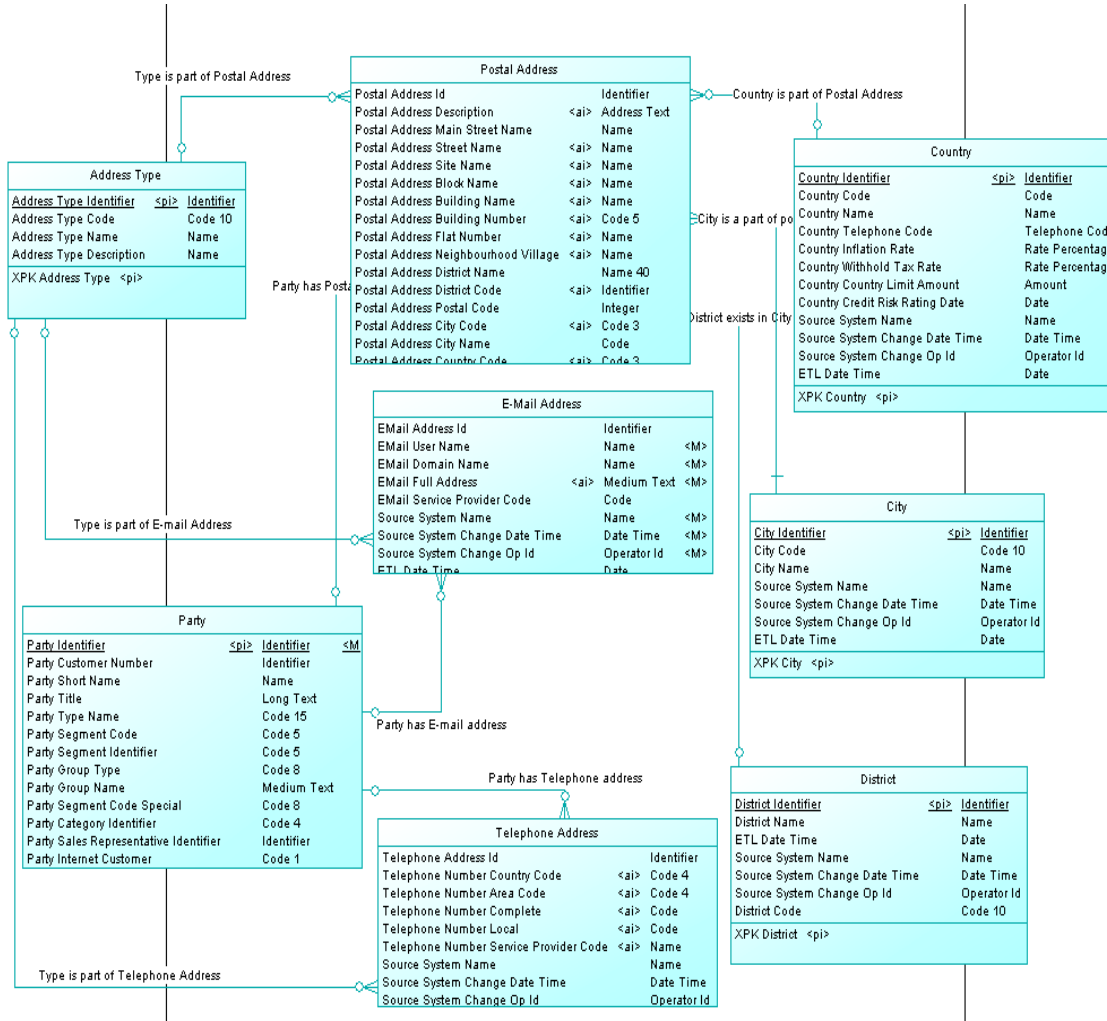
Aşağıda bankacılık sektörü için hazırladığım modelleme örnekleri verilmiştir.

Party Base(Müşteri Bilgileri - Genel) veri ambarı logical modeli Şekil 5.1 de gösterilmiştir.



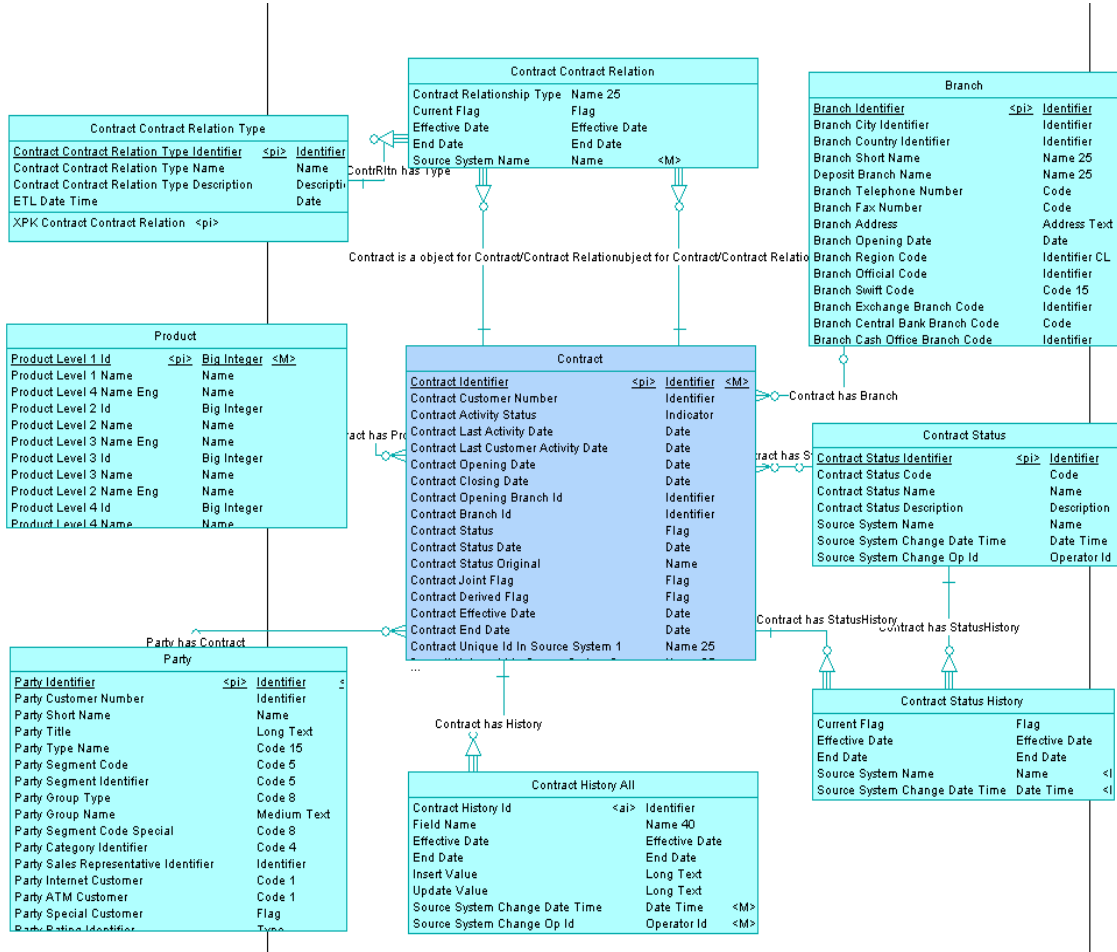
Şekil 5.1. Veri Ambarı Modelleme Diyagramı (Party Base)

Adress Base(Müşteri Adres Bilgileri) veri ambarı logical modeli Şekil 5.2 de gösterilmiştir.



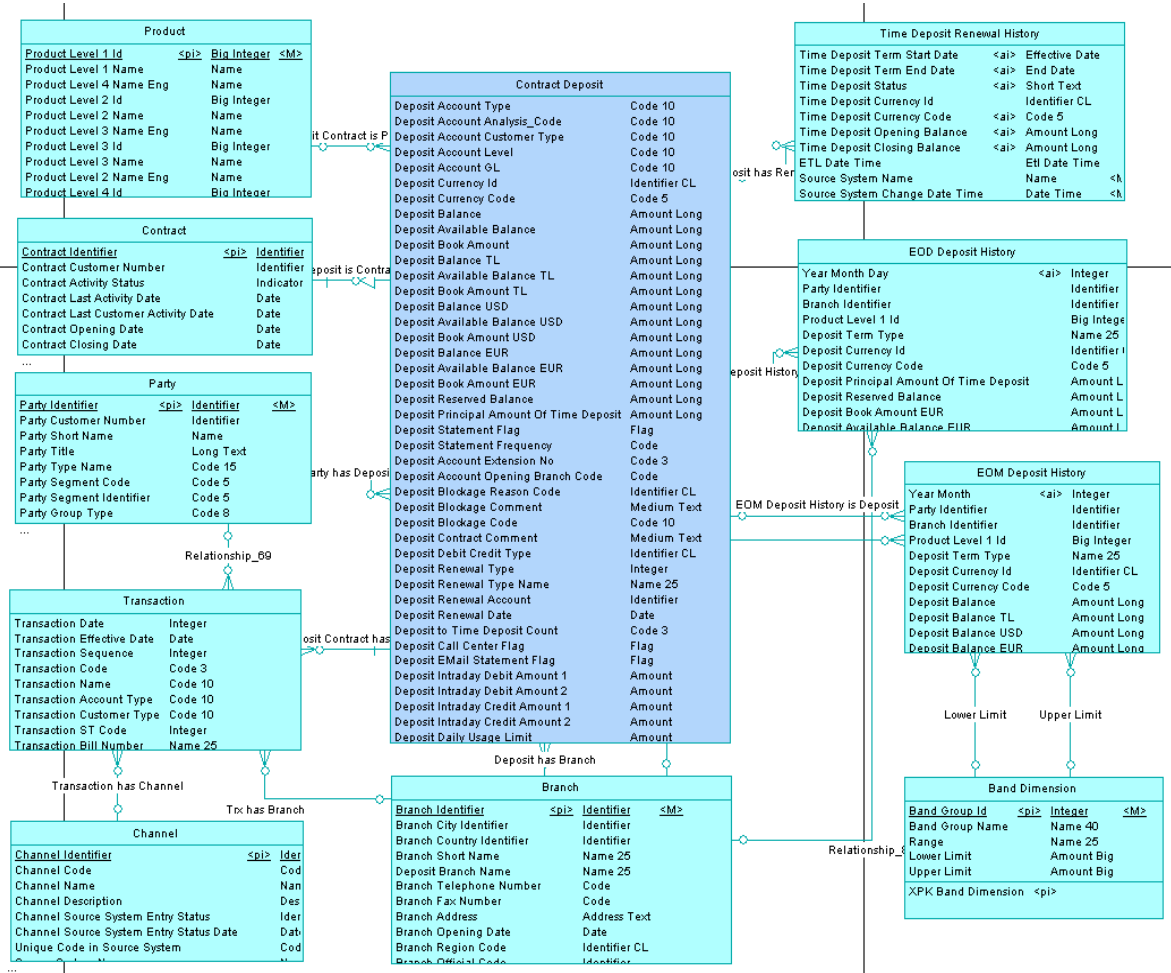
Şekil 5.2. Veri Ambarı Modelleme Diyagramı (Adress Base)

Contract Base(Müşterinin banka ile yaptığı ürün sözleşmeleri ortak bilgileri – Mevduat/Kredi/Yatırım gibi) veri ambarı logical modeli Şekil 5.3 de gösterilmiştir.



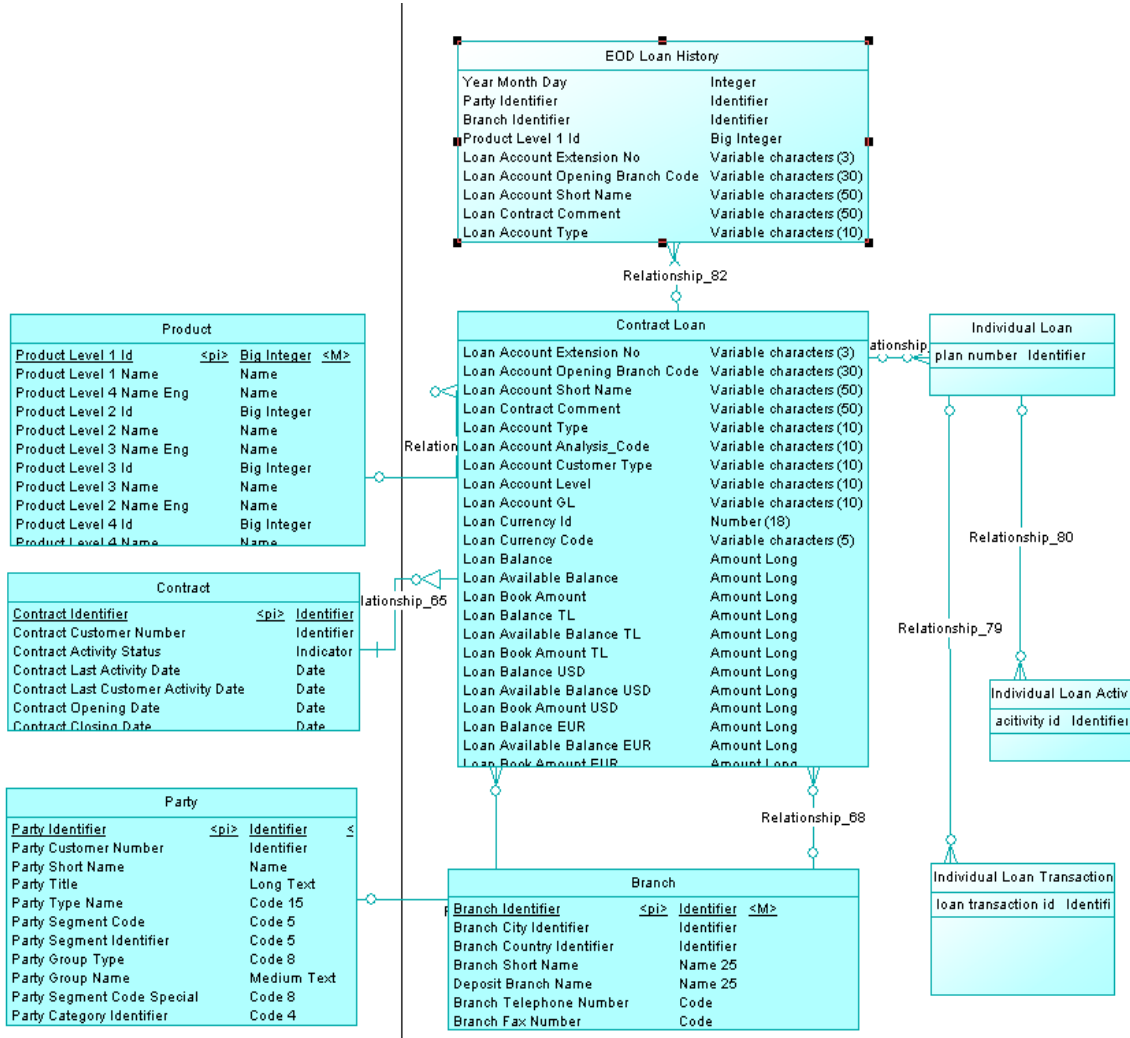
Şekil 5.3. Veri Ambarı Modelleme Diyagramı (Contract Base)

Contract Deposit(Müşterinin Mevduat Bilgileri – Ürün hiyerarşisi ile ilişkili) veri ambarı logical modeli Şekil 5.4 de gösterilmiştir.



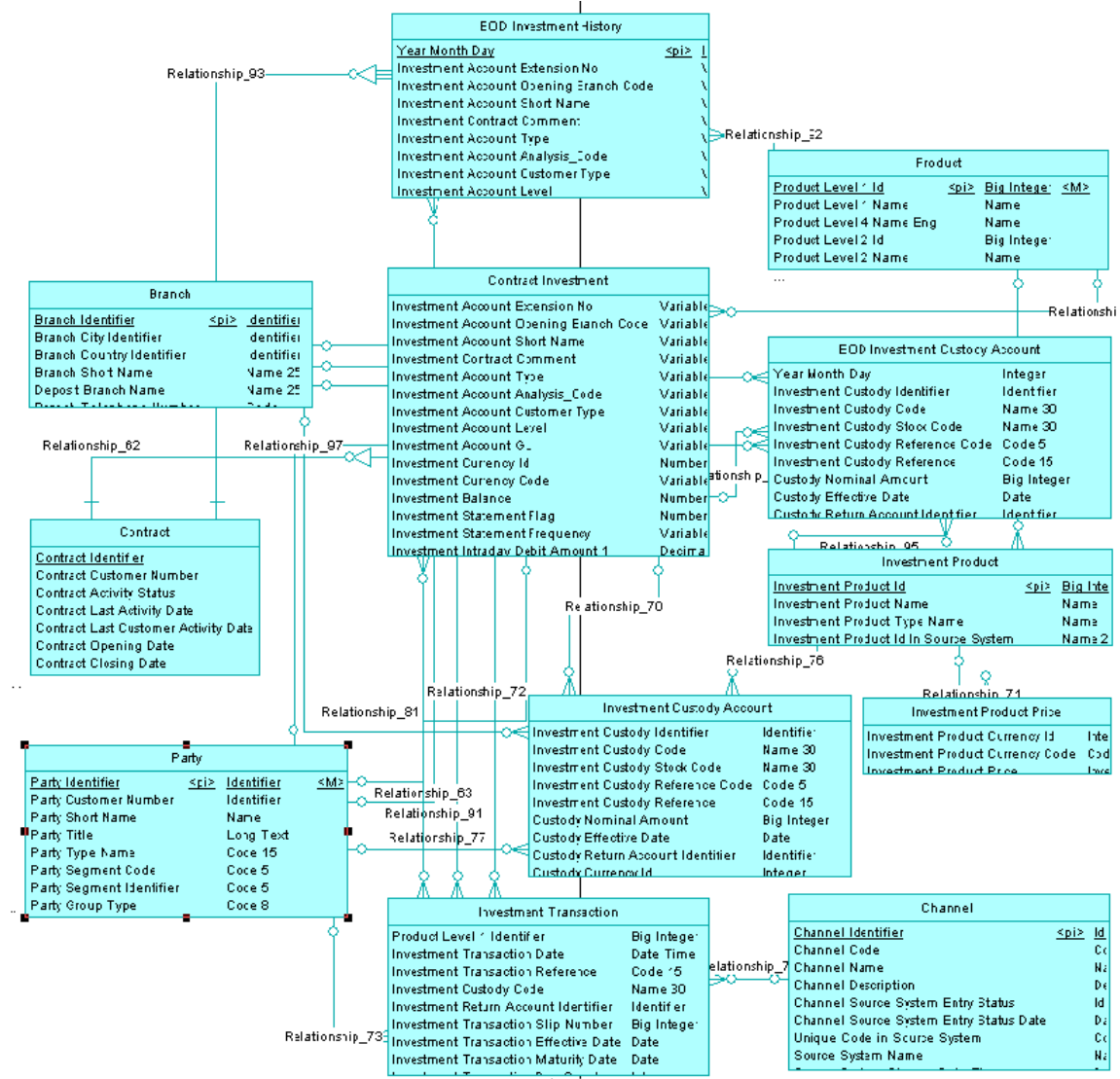
Şekil 5.4. Veri Ambarı Modelleme Diyagramı (Contract Deposit)

Contract Loan(Müşterinin Kredi Bilgileri – Ürün hiyerarşisi ile ilişkili) veri ambarı logical modeli Şekil 5.5 de gösterilmiştir.



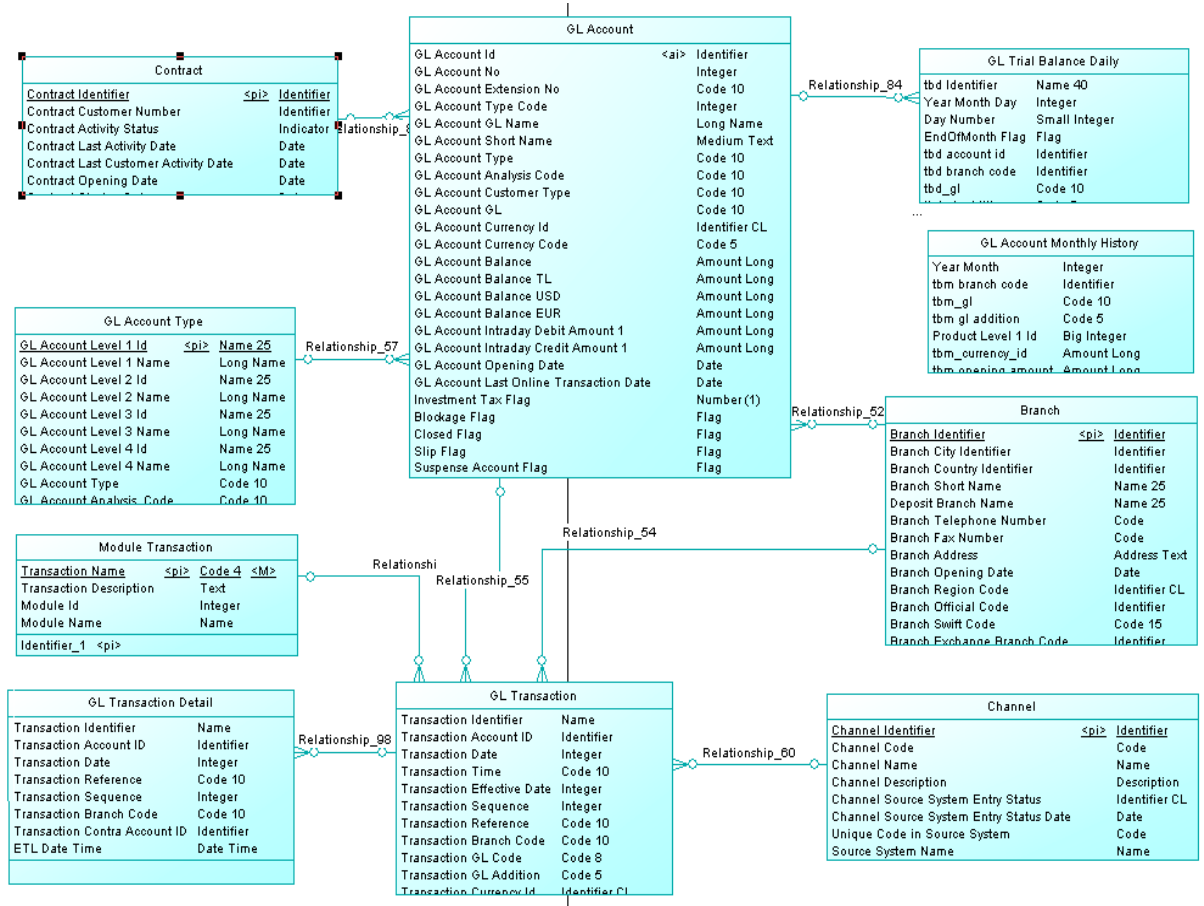
Şekil 5.5. Veri Ambarı Modelleme Diyagramı (Contract Loan)

Contract Investment(Müşterinin Yatırım Bilgileri (Repo/Fon/Hazine Bonosu/Eurobond) – Ürün hiyerarşisi ile ilişkili) veri ambarı logical modeli Şekil5.6 de gösterilmiştir.



Şekil 5.6. Veri Ambarı Modelleme Diyagramı (Contract Investment)

GL(General Ledger) Account (Transaction Bilgileri – Muhasebe/Contract/Kanal ile ilişkili) veri ambarı logical modeli Şekil5.7 de gösterilmiştir.



Şekil 5.7. Veri Ambarı Modelleme Diyagramı (GL Account)

5.6. BI Raporlama Platformu Modelleme

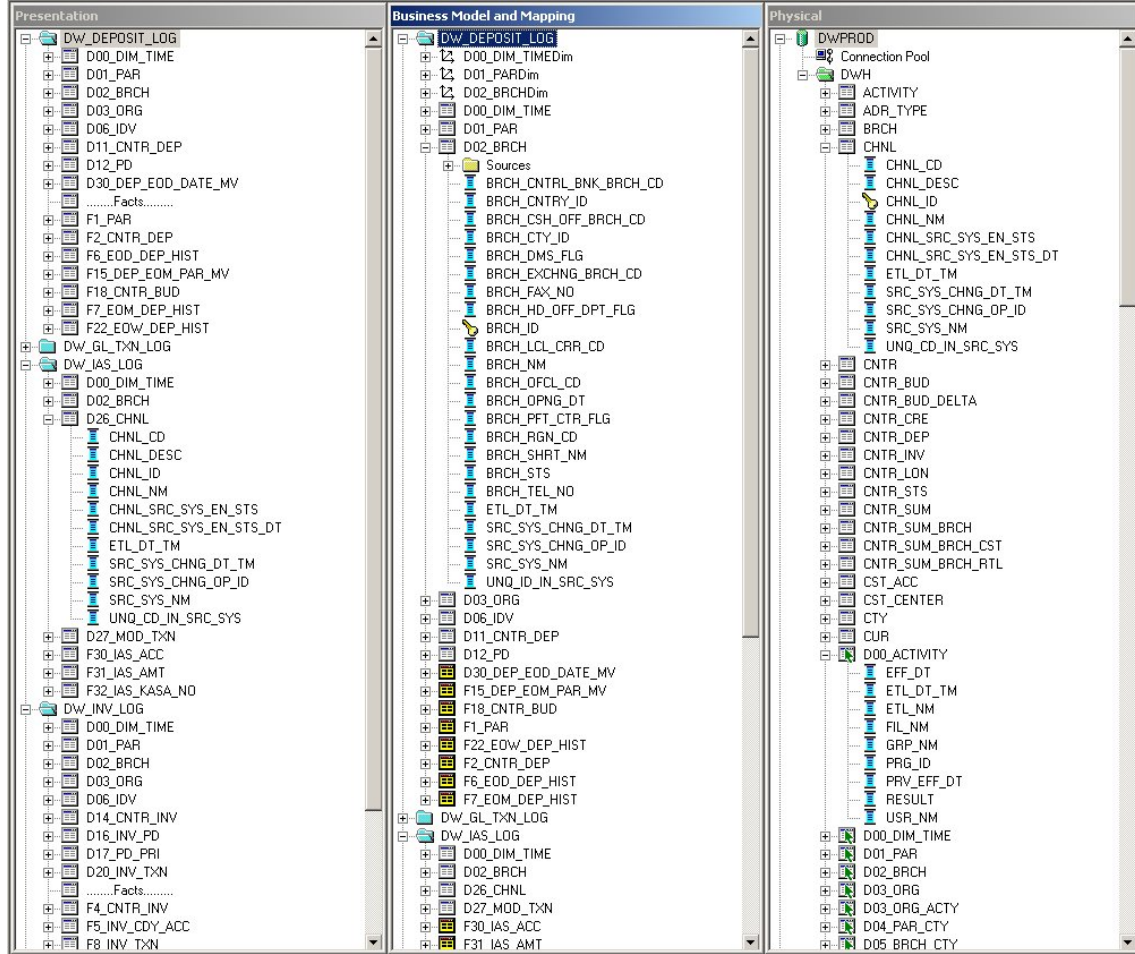
BI Raporlama Platformu modelleme ortamı üç layerden oluşmaktadır :

Physical layer, Veri ambarı üzerinde tanımlı tabloların fiziksel modellemesinin yapıldığı bölümdür.

Business Model and Mapping Layer, hiyerarşilerin tanımlandığı, özet alanların oluşturulabildiği, alanlar üzerine formülasyonların yapılabildiği ve fonksiyonların(sum, average gibi) tanımlanabildiği logical modelleme bölümüdür.

Presentation Layer, sürükle bırak yöntemi ile ansal raporlama hazırlanabilmesi için kullanıcıya açılacak alanların tanımlandığı ve modellendiği bölümdür.

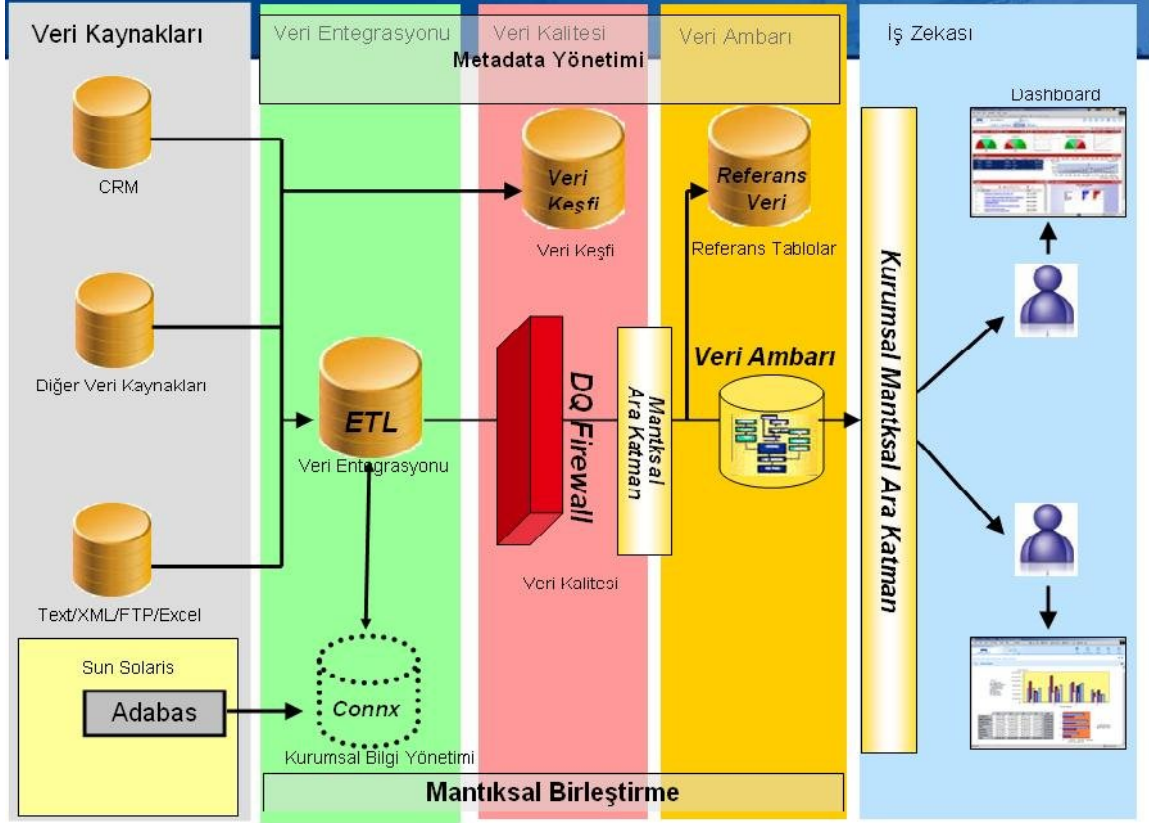
Bankacılık sektörü için hazırladığım BI raporlama platformu modelleme örneği Şekil 5.8 de gösterilmiştir.



Şekil 5.8. BI Raporlama Platformu Modelleme Örneği

5.7. ETL Süreci

Aşağıda bankacılık sektörü için hazırladığım bir ETL süreç örneği verilmiştir.



Şekil 5.9. BI ETL Süreci Örneği

Öncelikle en sol sütunda, kurum bünyesindeki veri kaynakları simgelenmektedir. Bu veri kaynakları temel olarak kurumun kullandığı veri tabanı (Oracle, SQL Server, DB2, Adabas vb.) olmakla birlikte, bugün gözönüne alınmayan ancak ileride gündeme gelebilecek ya da Text, XML, Excel gibi ekstra kaynaklar olabilir.

2. sütunda görüldüğü gibi ETL yazılım aracı tüm veri kaynaklarına ulaşarak bu verileri alıp işleyen ve uygun bir model ile tasarlanmış veri ambarına verileri yükleyen bir modüldür.

Veriler, 4. sütunda görüldüğü gibi veri ambarına aktarılmadan önce veri kalitesi operasyonları da gerçekleştirilecektir. Veriler ancak bir veri kalitesi kontrol duvarını geçebilirlerse (3. sütun) veri ambarına yüklenecektir.

Veri ambarı modelinin verimli ve hızlı raporlama, kolay kullanım ve kolay yönetim gibi özellikleri destekleyecek şekilde tasarlanmış olması gerekmektedir. Veri ambarına belli temizleme rutinleriyle aktarılmış olan kaliteli veri, bir süre sonra kaliteli bir referans veri kümesi oluşturacaktır ki bu veriler daha sonrasında tekrar operasyonel sistemlerde kullanılabilirler üzere gündeme gelebilecektir.

Son olarak veri ambarı üzerine yapılacak mantıksal tasarımın ardından kullanıcılar rapor ve dashboard'larına bir portal aracılığıyla erişebileceklerdir.

5.8. İş Zekası Projesi Yönetim Önerileri

İş Zekası projeleri çok fazla katmandan oluştuğu için alt yapısal anlamda karmaşık projelerdir. Business anlamda çok fazla konu içerirler. İhtiyaç analizinin yapılması, business modelin çıkarılması ve mevcut yapı ile map edilmesi uzun zaman alıcı işlemlerdir. Uzun soluklu projelerin yönetimi(Özellikle alt yapı oluşumu uzun süre alan), kısa süre içerisinde projeden beklentisi olan yönetim kadrolarına ve iş kollarına isteklerin verilememesi nedeniyle motivasyonun bozulması ile zorlaşmaktadır.

Bu nedenle iş zekası projelerinde dikkat edilmesi gereken noktalar ve öneriler aşağıda belirtilmiştir. (Bankacılık sektörü baz alınmıştır. Genele uygulanabilir.)

- ✓ Üst Yönetim farkındalığı ve desteği
 - Üst yönetime proje anlatılmalı ve desteği alınmalıdır. Proje sponsoru mutlaka Genel Müdür Yardımcısı düzeyinde olmalıdır.
- ✓ İş Birimleri ile uyumlu işbirliği
 - İş birimlerinin projeye aktif olarak katılımcı olmaları sağlanmalıdır.
- ✓ Proje desteğine sürekliliğin sağlanması ve proje kazanımlarının uygunluğunun temini için fazlandırma
 - Proje motivasyonunun düşmemesi için proje fazlandırılmalı ve anlamlı süre içerisinde proje beklentisi olan yönetim kadroları ve iş kollarının beklentileri karşılanmalıdır.
- ✓ Yönetişim yapısının kurulması
 - Kurulan yapının düzenli olarak çalışması ve devamlılığının sağlanması için yönetişim yapısı kurulmalıdır.
- ✓ Proje ekiplerinin oluşturulması
- ✓ İş zekası projesi için kullanacağınız platform ve ETL tool seçimi için ön eleme yapılmalı ve ön elemeyi geçen firmalardan tanıtıcı sunum alınmalıdır.
- ✓ Tool seçimi sürecinde, Tool'un özelliklerinin incelenmesi ve birbirleri ile karşılaştırılabilmesi için Kavram Kanıtlama (PoC) çalışması yapılmalıdır. (Hazırlanmış senaryo kullanılarak)
- ✓ Çerçevesi iyi çizilmiş (Beklentilerin tam ve net olarak anlatılması) bir RFP (Request for Proposal) hazırlanmalı ve firmalardan teklifler bu RFP üzerinden alınmalıdır.
- ✓ Sektörel ve teknolojik gelişmelere ve değişimlere cevap verebilecek esnek bir model tasarımı yapılmalıdır. (Veri Ambarı ve BI Model)

- ✓ Donanım büyüklük tespiti (Disk size estimation) yapılmalıdır.
- ✓ Arşiv yönetimi planlanmalıdır.
- ✓ History datasının ne kadar geriden başlanarak tutulacağı ve yönetimi planlanmalı ve tasarlanmalıdır.

5.9. İş Zekası Platform ve ETL Tool Seçimi

İş Zekası kurulumunun en önemli konularından bir tanesi kullanılacak olan platformun seçilmesidir. Doğru seçim proje için gerçekten çok önemlidir. Bu yüzden üzerinde detaylı çalışılmalıdır. Mutlaka örnek bir senaryo hazırlanmalı ve tercih edilmesi düşünülen firmalarla bu senaryo üzerinde çalışılmalıdır. Firmanın yaklaşımı, senaryoyu gerçekleştirimi detaylı bir şekilde izlenmeli ve değerlendirilmelidir.

İş Zekası Platformu ve ETL aracını değerlendirebilmek için detaylı bir soru seti hazırlanmalıdır. Firmalara bu soru setleri gönderilmeli ve platformlarının gerçekleştirme puanlamaları istenmelidir.

Ayrıca bu konuda çerçevesi çok iyi çizilmiş bir RFP(Request for Proposal) hazırlanmalı ve firmalardan teklifleri bu çerçevede hazırlamaları istenmelidir.

Aşağıda İş Zekası Platformu ve ETL Tool seçimi için hazırladığım soru kümeleri verilmiştir.

İş Zekası Platformu:

- Teklif edilen iş zekası yönetim sistemi, çok boyutlu veri modelleme ve işleme (OLAP), veri madenciliği, standart ve anlık raporlama hizmetleri için gerekli servis ve araçları içermelidir.
- Aşağıda detaylandırılan analiz, veri madenciliği, raporlama ve veri aktarım araçları, aynı arayüzlerle yönetilebilmeli, böylece uygulama geliştirici ve sistem yöneticilerinin performansı artırılmalıdır.

Analiz, Raporlama ve Sistemsel Özellikler:

- Son kullanıcıların teknik alt yapıyı bilmeden sorgulama yapabilecekleri ara katman, görsel bir ortamda sürükle/bırak özellikleri desteklenerek geliştirilebilmelidir.
- Kullanıcıların sorgu yaratırken kullanacağı sütunlar, veritabanındaki tablolara ait alanlar olabileceği gibi, daha karmaşık formüller de olabilir. Bu sırada farklı tablolardaki sütunlar arasında toplam/çarpım gibi işlemler yapılabileceği gibi, veritabanının özgün (native) olarak desteklediği fonksiyonlar da kullanılabilir.
- Ara katman üzerinde bir değişiklik yapıldıktan sonra bu ara katmanı kullanarak daha önceden hazırlanmış raporlar etkilenmemelidir.

- Ara katman tasarımı sırasında, Geliştirme/Test/Üretim gibi farklı aşamalar desteklenmelidir. Geliştirme aşamasındaki ara katmanlara kullanıcılar ulaşamamalı ve gerekli testler yapıldıktan sonra ara katmanlar sorunsuz olarak üretime aktarılabilmelidir.
- Geliştirme ortamında tablolar arasında basit ilişkiler kurulabileceği gibi, birden fazla sütunu kapsayan daha karmaşık Join ilişkileri kurulabilmelidir.
- Geliştirme ortamı self-join özelliğini desteklemelidir.
- Kullanıcılar önceden tanımlanmış bir nesneyi seçtiklerinde, bir prompt ekranında kısıtlama girmeleri zorunlu kılınabilmeli. (Örneğin Masraf alanı bir kullanıcı tarafından rapora alındığında, mutlaka başlangıç ve bitiş tarihlerinin girilmesi aksi takdirde raporun çalıştırılmaması gibi).
- Ara katman tasarımı sırasında veritabanındaki tablo ve view'lar kullanılabilmesi gibi, gerekirse dinamik view'ların yaratılması mümkün olmalıdır. Dinamik view'ların yaratılması işleminin veritabanında yapılmasına gerek olmamalıdır.
- Bir tabloya kısıtlama girilebilmelidir. Bu sayede bir kullanıcı o tablodan hangi alanı seçerse seçsin, otomatik olarak ilgili kısıtlamanın seçilmesi gerçekleştirilebilmelidir.
- Veritabanında yaratılmış özet tablolar otomatik olarak algılanmalı ve bir sorguya ait SQL cümleciğinin oluşturulması sırasında gerekmiyorsa detay tablolara gidilmemelidir. Altyapı tanımlandıktan sonra bu işlem otomatik olarak gerçekleştirilmelidir.
- Veritabanında tanımlı olan ve bir result set döndüren Stored Procedure'lar ara katman seviyesinde bir tablo gibi kullanılabilmelidir.
- List of Values tanımlanırken, eğer tanımlı bir indeks varsa otomatik olarak algılanmalıdır.
- Ara katmana ait bilgiler, merkezi bir alanda (repository) saklanabilmelidir.
- Repository içinde raporlar saklanabileceği gibi Word, Excel, Powerpoint ya da PDF gibi farklı formattaki dökümanlar da saklanabilmelidir. Bu raporlar da son kullanıcılar tarafından görüntülenebilmelidir.
- Raporlama sisteminde kullanıcı ve gruplar tanımlanabilmeli ve kullanıcı ya da grup bazında güvenlikle ilgili haklar verilebilmelidir.
- Güvenlik modülünde, hangi kullanıcı ya da grubun hangi sütunlara ulaşabileceği tanımlanabilmelidir.

- Satır bazında güvenlik desteklenmelidir (Örneğin bir rapor farklı şubelerden açıldığında, her şubenin kendisine ait olan bilgileri görmesi gibi..)
- Hangi raporlara kimlerin erişebileceği ya da kimlerin erişemeyeceği tanımlanabilmelidir.
- Hangi sütuna hangi kullanıcıların erişebileceği ya da erişemeyeceği tanımlanabilmelidir.
- Son kullanıcı modülleri ister istemci sunucu mimaride isterse Web ortamında kullanılmalı, tüm güvenlik ayarları tek bir merkezden aynı şekilde kontrol edilebilmelidir.
- Güvenlik ayarları nesne bazında yapılabilir. Örneğin A kişisi bir X raporunu refresh edebilecekken bir Y raporunu yenileme yetkisi alınabilmelidir.
- Active Directory, NTLM gibi güvenlik standartları desteklenmelidir.
- 128 bit SSL güvenlik standardı desteklenmelidir.
- Service Oriented Architecture desteklenmelidir.
- Raporlar ön bellekte(cache) tutulabilmelidir. Ön belleğe alınan bir rapor, parametrik olarak verilen süre içinde tekrar tekrar açıldığında veritabanına gitmeye gerek kalmadan bu rapor ön bellekten açılabilir. Bu işlem kullanıcı yenileme düğmesine basarsa ya da bir kullanıcı tarafından ön belleğe alınan bir rapor başka bir kullanıcı tarafından açılabilir gerçekleştirilebilir.
- Ön Belleğe alma işlemi sunucu bazında gerçekleştirilebileceği gibi rapor bazında da yapılabilir. Bir X raporu ön belleğe alınırken bir Y raporu gerek yoksa ön belleğe alınmadan kullanılabilir.
- Sunucu prosesleri hem yatay hem de dikey genişlemeye izin verecek şekilde ölçeklenebilir. Örneğin bir cluster yapısında bir prosesin hangi makinede çalışacağına ve aynı anda kaç instance'ının olacağı parametrik olarak belirlenebilir.
- Raporlama platformu, son kullanıcıların teknik alt yapıyı bilmeden sorgulama yapabilecekleri bir ortam sunmalıdır.
- Son kullanıcı raporlama yazılımı DHTML raporlama paneli, Java Aplet ya da rapor sihirbazları gibi farklı opsiyonlar sunarak farklı profildeki kullanıcıları desteklemelidir.
- Raporların tasarımı raporlama platformunda sürükle ve bırak özelliklerini kullanarak yapılabilir.

- Raporlar son kullanıcı bilgisayarına herhangi bir ürün kurmadan yaratılabilmelidir.
- Son kullanıcı yazılımı, farklı ilişkisel veritabanlarına özgün (native) olarak erişebilmelidir.
- Son kullanıcı yazılımı, MS OLAP/Analysis Server, DB2 OLAP Server, SAP BW gibi çok boyutlu (multidimensional) veritabanlarına özgün(native) olarak erişebilmelidir.
- Bir döküman içinde, birden fazla sorgu yaratılabilmelidir.
- Aynı döküman içinde, farklı verikaynakları sorgulanabilmelidir. Örneğin bir sorgu Sybase/Oracle/MS SQL Server/DB2 gibi ilişkisel veritabanından veri getirirken aynı döküman içindeki ikinci bir sorgu da başka bir ilişkisel veri kaynağından ya da bir diğer sorgu da herhangi bir başka OLAP veritabanından gelebilmez. Bu sorguların sayısında bir sınırlama olmamalıdır.
- Farklı verikaynaklarından sorgulanan bilgiler, arada mantıksal bir ilişki varsa rapor üzerinde aynı tabloda birleştirilebilmelidir. Örneğin rapor üzerindeki bir tabloda bir sütun Sybase/Oracle/MS SQL Server/DB2 gibi bir veritabanından, 2. sütun başka bir ilişkisel verikaynağından bir diğer sütun ise MS Analysis Server'dan gelebilmelidir. Bu sütunlar arasında işlem yapılabilmelidir.
- Hazırlanan rapor sayfalarında aynı anda birden fazla bilgi blokları oluşturulabilmelidir.
- Hazırlanan raporlarda aynı anda hem tablo bilgileri hem de grafikler görüntülenebilmelidir.
- Bir döküman içinde birden fazla worksheet (sayfa) oluşturulabilmelidir.
- Bilgiler verikaynaklarından çekildikten sonra rapor tasarımı yaparken verikaynaklarına yeniden erişmeye gerek kalmamalıdır.
- Veriler rapor üzerinde görüntüledikten sonra, tüm hesaplama işlemleri bir daha verikaynağına erişmeye gerek kalmadan yapılabilmelidir.
- Rapor bilgileri üzerinde character, string, date-time, RunningSum() gibi fonksiyonlar desteklenmelidir.
- Raporlar üzerinde formüller ve değişkenler tanımlanabilmelidir. Bu değişkenler tekrar tekrar kullanılabilme özelliğine sahip olmalıdır.
- Raporlar Excel formatında kaydedilebilmelidir. Bu şekilde kaydedilen raporlar, formatları kaybolmadan, eğer varsa grafiklerle birlikte Excel'e aktarılabilir.
- Hazırlanan dökümanlar, Web ortamından görüntülenebilmelidir.

- Dökümanlar Web ortamında görüntülendikten sonra istenirse güncellenebilmelidir. Güncelleme yaparken eğer prompt tanımları yapıldıysa kullanıcılar güncellemeyi bu prompt bilgilerini girdikten sonra yapabilmelidir.
- Rapor kategorileri oluşturulabilmelidir. Raporlar gruplanarak farklı kategorilere atanabilmeli ve belli bir kategori seçildiğinde sadece o kategoriye ait raporlar listelenebilmelidir.
- Raporlar hazırlandıktan sonra başka kullanıcılara gönderilebilmelidir.
- OLAP analizi ilişkisel veritabanlarına karşı yapılabilmelidir.
- Çok boyutlu veritabanlarına karşı OLAP analizi yapılabilmelidir.
- OLAP veritabanları sorgulandıktan sonra, ranking, nested ranking, filtering, nested filtering, family fonksiyonları gibi özellikler desteklenmelidir.
- OLAP veritabanlarından bilgi getirildikten sonra yeni sütunlar yaratılabilir ve bunlar üzerinde formüller tanımlanabilmelidir.
- OLAP veritabanlarından ilişkisel veritabanlarına drill-through özelliği desteklenmelidir. Örneğin MS OLAP Server üzerinde drill down yapılırken gerekli tanımlar yapıldıysa drill-through yöntemiyle Sybase gibi bir veritabanına karşı hazırlanmış bir rapora dallanılabilir. Bu sırada OLAP veritabanı üzerinde hangi hücre üzerinde drill-through yapılıyorsa o hücre bilgisi, ilişkisel veritabanına karşı hazırlanan rapora parametre olarak geçirilebilir.
- Report Linking (Document Linking/Drill Through) fonksiyonlitesiyle bir rapordan diğer bir rapora geçiş sağlanabilir. Bu sırada ikinci rapora otomatik parametre gönderilmesi sağlanabilir.
- Rapor yaratabilen kullanıcılar Report Linking özelliğini bilgi işlem departmanına ihtiyaç duymadan kendileri oluşturabilir.
- Web üzerinde yaratılan raporlar sunucu üzerinde kaydedilebileceği gibi gerektiğinde yetkisi olan kullanıcılar tarafından kendi makinelerine de kaydedilebilir.
- Kendi makinesi üzerine rapor kaydeden bir lisanslı kullanıcı yetkisi dahilinde bu raporu sunucuya bağlanmadan da açabilir ve üzerinde analiz yapabilir.
- Kendi makinesi üzerinde bulunan rapor üzerinde değişiklik yapan kullanıcı tekrar sunucuya bağlanarak raporu güncellediği zaman (refresh), hangi satırların silindiğini, hangi satırların değiştiğini ya da eklendiğini görsel olarak takip edebilir.

- MS Analysis, DB2 OLAP Server ya da Oracle OLAP Server kullanılarak hazırlanan bir OLAP raporundan, ilişkisel veritabanına karşı hazırlanmış bir rapora Report Linking fonksiyonallitesi desteklenmelidir.
- Web arayüzü .NET ya da JSP desteği sunmalı ve ekranlar bu yöntemlerle değiştirilebilmelidir.
- Promptlara öndeğer atanabilmelidir.
- Subquery'ler yaratılabilmelidir.
- Raporlar görüntülenirken sayfa sayfa görüntülenebilmeli ve istenildiği zaman sayfa numarası yazılarak belli bir sayfaya direk olarak gidilebilmelidir.
- Raporlar, istenildiği kaydedilebilmelidir. Kaydetme sırasında sadece raporun formatı kaydedilebileceği gibi istenirse rapor formatıyla birlikte verilerin kendisi de saklanabilmelidir.
- Raporlar üzerinde filtreler konulabilmelidir. Bu sırada tekrar veritabanına gitmeye gerek kalmamalıdır
- Raporlar üzerinde sıralama işlemi yapılabilmelidir. Bu sırada tekrar veritabanına gitmeye gerek kalmamalıdır.
- Web ortamında değişken ve formüller tanımlanabilmelidir. Bu değişkenler aynı rapor içinde tekrar tekrar kullanılabilir.
- Raporların periyodik ve otomatik olarak güncellenebilmesi gerekmektedir.
- Periyodik ve otomatik olarak güncellenen raporlar Excel ya da PDF gibi farklı formatlarda saklanabilmelidir.
- Periyodik olarak güncellenen raporlar, raporlama sistemi içinde başka kullanıcılara gönderilebilmelidir.
- Periyodik olarak güncellenmiş raporlar, farklı kullanıcılar tarafından açıldığı zaman her kullanıcı yetkili olduğu sütun ve satırları görmelidir.
- Raporlar periyodik olarak güncellendikten sonra organizasyon içindeki diğer son kullanıcılara email'e attach edilerek gönderilebilmelidir.
- Raporlar emaille organizasyon içi ya da dışındaki kullanıcılara otomatik olarak gönderilebilmelidir. Bu sırada emaillerin kimlere ne zaman gönderileceği bir administration yazılımıyla belirlenebilmelidir.
- Raporlar direk BusinessObjects formatında olabileceği gibi, Excel ya da PDF gibi formatlarda da otomatik olarak gönderilebilmelidir.

- Periyodik olarak güncellenen raporlar istenildiği takdirde otomatik olarak versiyonlanabilmeli ve daha önce kaydedilmiş bu eski versiyonlara erişebilmelidir. Eski tarihli bir rapor ve son hali bir arada görüntülenebilmelidir.
- 3rd Party Scheduling yazılımlarıyla entegre olabilmelidir.
- Custom Calendar tanımlayarak periyodik olmayacak şekilde, özel tarihlerde raporların güncellenmesi sağlanabilmelidir.
- Sisteme bağlanan kullanıcıların kim oldukları, hangi saatlerde bağlandıkları, hangi raporları kullandıkları, hangi SQL cümleciklerini çalıştırdıkları gibi bilgilere erişilebilmelidir.
- Sistemde uzun süre kullanılmayan ve hiç açılmayan raporların listesi alınabilmelidir.
- Sistemde en çok kullanılan raporların listesi alınabilmelidir.
- En fazla ulaşılan nesnelere ya da tablo field'larının listesi alınabilmelidir.
- Geriye doğru analiz yapılarak veritabanında bir field değiştirildiğinde bunun hangi raporları etkileyeceği bulunabilmelidir.
- Tüm yukarıdaki sorgular verilirken tarih aralığı gibi kısıtlamalar verilebilmelidir.
- Oluşturulan raporlara ait tartışma grupları yaratılabilmeli. Bu tartışma gruplarında rapora ait bir mesaj ilgili kişi ya da gruplara gönderilebilmeli ve bu mesaja ait yanıtlar bir ağaç yapısında tarihsel olarak izlenebilmelidir.
- Oluşturulan raporların hangi amaçla yaratıldığı, içinde hangi sorguların çalıştırıldığı, her sorguda hangi alanların sorguya dahil edildiği ve hangi prompt kriterlerinin kullanıldığı gibi bilgiler bir ansiklopedi tarzında takip edilebilmelidir.
- Bir raporun hangi sorulara yanıt verdiği, içinde kullanılan terimlerin neler olduğu gibi bilgiler ansiklopedik olarak saklanabilmeli, gerektiğinde soru bazında search işlemi yapılabilmelidir.
- Microsoft Office desteği olmalıdır. Oluşturulan raporlar Word, Excel, Powerpoint gibi yazılım araçları içinden güncellenebilmeli ve sonuçlar direk olarak Word, Excel ya da Powerpoint içine alınabilmelidir. Bu sırada gerekirse prompt kriterleri girilebilmelidir.
- Web üzerinde Ad-hoc sorgular oluştururken SQL cümlecikleri minus, intersect, union gibi operatörleri desteklemelidir.
- Sorgulara kısıtlama koyarken iş terimleri birbiriyle karşılaştırılabilmelidir (object-to-object Comparison)
- Oluşturulan SQL cümlecikleri, eğer ilgili kullanıcının yetkisi varsa değiştirilebilmelidir.

- Hem raporda hem de SQL cümleciklerinde ranking fonksiyonları desteklenmelidir.
- Custom Sort özelliği olmalıdır.
- Rapor bazında ranking fonksiyonu desteklenmelidir.
- Raporlarda alert (conditional formatting) özelliği desteklenmelidir.
- Sistemde dashboard ekranları tanımlanabilmelidir.
- Dashboard ekranlarında bir programlama yapmaya gerek kalmadan kritik başarı faktörleri speedometer (hız göstergeleri), termometre ya da trafik işaretleri gibi simgelerle gösterilebilmelidir.
- Kritik başarı faktörlerinin gösterildiği hız göstergeleri, termometre ya da trafik işaretleri gibi simgeler üzerinde trend analizi yapılabilir.
- Kritik başarı faktörlerinin gösterildiği hız göstergeleri kullanıcılar tarafından etkileşimli olarak değiştirilebilmelidir. Kullanıcı zaman kriterini değiştirerek bu göstergeleri ekranda etkileşimli olarak güncelleyebilmelidir.
- Üzerinde hız göstergesi, termometre gibi bileşenler bulunan Dashboard ekranları PDF, Flash, Powerpoint, Word, Outlook gibi ortamlara kaydedilebilmelidir.
- PDF, Powerpoint, Word, Outlook, Flash gibi ortamlara kaydedilen dashboard ekranlarında kullanıcılar etkileşimli analiz, what-if analizi yapabilmeli ve gerektiğinde oluşturulan dashboard'ları bu ortamlardan güncelleyebilmelidir.
- Hız göstergeleri, termometre ya da trafik işaretleri gibi simgelerden raporlara drill down yapılabilir, bu sırada parametre geçişleri otomatik olarak sağlanabilir.
- Birbiriyle ilişkili kritik başarı faktörleri bir ağaç şeklinde son kullanıcıya sunulabilir.
- Raporlara mesajlar eklenebilir ve tartışma grupları oluşturulabilir.
- Alerter (Exception Highlighting) yapılırken object-to-object comparison yapılabilir.
- Alerter tanımlarken Alert condition'una uyan kompleks mantıksal ifadeler kullanılabilir.
- Alerter tanımlarıyla, renklerin, yazı karakterlerinin değiştirilmesi mümkün olabileceği gibi, hücre seviyesine başka bir formül ya da değişken değerinin atanması da mümkün olmalı.
- Minus, Intersect ve Union gibi bileşik sorgular, son kullanıcıların SQL bilgisine ihtiyaç duymadan yaratılabilmesine olanak tanımalıdır.
- Otomatik olarak oluşturulan SQL cümlecikleri, eğer kullanıcının yetkisi varsa değiştirilebilmelidir.

- Prompt kriterlerinin hangi sırayla görüntüleneceği programlama yapmadan, görsel olarak belirlenebilmelidir.
- Custom Sort yapılabilmesi.
- Custom Sort yapılırken, o an raporda olmayan, ancak daha sonraki bir yenileme sırasında gelebilecek değerler de sort kriteri olarak verilebilmesi.
- Hazırlanan raporlarda ranking yapılabilmesi.
- Ranking yapılırken, rank yapılacak rapor tablosunda istenilen sütun üzerine ranking uygulanabilmesi.
- Rank fonksiyonu, bir sütundaki tüm değerleri sıralamak amacıyla da kullanılabilmesi.
- Ürün, son kullanıcıların ister ilişkisel veritabanı üzerinde isterse rapor bazında ranking yapmasına olanak tanımalıdır.
- Oluşturulan ara katman, veritabanları CWM(Common Warehouse Meta-Model) metadata standardını desteklemeli.
- Raporlama yazılımı, Oracle Warehouse builder, IBM Data Warehouse Center, ya da IBM Cubeview gibi yazılımlarla metadata exchange yapabilmeli.
- CWMstandardında metadata exchange yapılabilmesi.
- Raporlama yazılımı Oracle OLAP küplerini otomatik olarak okuyarak kendi ara katmanını oluşturabilmeli. Bu işlem görsel olarak yapılabilmelidir.
- Raporlama yazılım aracı, RunningSum(), RunningProduct(), Standard Deviation gibi istatistik fonksiyonları desteklemelidir.
- Oluşturulan raporlardaki bloklar, Excel, Powerpoint, Word gibi ortamlardan refresh edilebilmelidir. Örneğin bir powerpoint prezentasyonu içinden refresh düğmesine basan bir kullanıcı, eğer varsa parametre girişlerini yapmalı ve raporun son hali prezentasyon içine getirilebilmelidir.
- X arakatmanı kullanılarak hazırlanan bir rapor, istenildiğinde Y arakatmanına bağlanabilmelidir. Bu sırada eğer arakatmanlarda farklı isimler kullanılıyorsa, X arakatmanından oluşturulan rapordaki sütunlar, Y arakatmanındaki sütunlarla eşleştirilebilmelidir.
- Sorgu bazında kaç satır bilginin getirilebileceği tanımlanabilmelidir.
- Bir sorgunun maksimum kaç dakika çalışacağı belirtilebilmelidir. Bu süreden uzun süren sorgular otomatik olarak kesilebilmelidir.

- Rapor bazında oluşturulan sorguların, başka kullanıcılar tarafından değiştirilememesi hakkı verilebilmeli. Bu durumda sadece raporu oluşturan kişi, sorguyu değiştirebilecektir.
- Raporda oluşturulan prompt'ların ekranda görüntülenme sırası, görsel olarak değiştirilebilmelidir.
- Bir sorgu çalıştırılmadan ve veri seti çekilmeden rapor tasarımı yapılabilirdir.
- Bir raporda birden fazla sorgu varsa, bunlardan herhangi biri çalıştırılabilirdi ve sadece 1 tanesini çalıştırmak için hepsini çalıştırmaya gerek olmamalıdır.
- Bir raporda birden fazla sorgu varsa bu sorguların hangi sırada çalışabileceği belirlenebilmelidir.
- Drill Down/Up işlemi yapılırken, bu fonksiyonalitye istenildiğinde istemci bilgisayarda, istenildiğinde sunucu veritabanı tarafında (Query Drill) gerçekleştirilebilmeli.
- Distinct count, Rank, Standard Deviation, Variance ve Running Aggregate gibi hesaplanmış sütunlar için yapılan Drill Down/Up işlemi doğru sonuç üretmelidir (Query Drill).
- Bir raporda, birden fazla block oluşturulduğunda bunlardan birisini sayfa üzerinde drag&drop ile hareket ettirirken/taşırken, istenildiğinde diğer block'lar da relative olarak otomatik biçimde taşınmalıdır.
- Bir gruba ya da kişiye rapor içinde yapılabilecek aksiyonlar bazında yetki verilebilmeli. Örneğin, bir A kişisi, X raporu üzerinde SQL'i edit edebilme veya raporu refresh edebilme gibi bir yetkiye sahip olurken, bu haklar aynı kişi için , Y raporu üzerinden alınabilmeli ve böylece bu örnekte Y Raporunun SQL'i değiştirilememeli veya Y raporu yenilenememelidir.

Performans Yönetimi Sistemi :

- Performans yönetimi sistemi, temel performans göstergeleri (key performance indicators – KPI) kendi üzerinde tanımlayabilmelidir.
- Temel performans göstergeleri “ölçütler” (metric) halinde biraraya getirilebilmelidir.
- Temel performans göstergeleri veya ölçütlerle ilgili detay analizler, ek rapor ve grafiklerle detaylı olarak açıklanabilmelidir.
- Temel performans göstergeleri, hedef ve fiili değerleri verebilmeli; bu hedef ve fiili değerler üzerinden trafik ışıkları, v.b. göstergeler tanımlanabilmelidir. Performans yönetim sistemine yeni göstergeler, kurum tarafından kolaylıkla ve ek programlama gerektirmeden tanımlanabilmelidir.

- Kurum stratejilerinin şirket aksiyonları, departmanlar ve temel performans göstergeleri ile uyumlandırılabilmesi için strateji haritası geliştirme imkânı bulunmalıdır. Kurum strateji haritasında yer alan inisiyatifler, veritabanlarında yer alan fiili ve hedef değerlerle ilişkilendirilebilmelidir.
- Performans yönetimine dair hedef değerler, bütçe, operasyonel plan, stratejik plan gibi farklı planlardan gelebilmelidir. Bu hedef değerler aynı sistem içerisinde tanımlanabilmelidir.
- Performans Yönetim Sistemi, “hedef” değerlerini, planlama uygulamasından gelen bütçe veya tahmin değerlerinden alabilmelidir.
- Kişi, departman, bölüm, kurum geneli gibi farklı seviyelerde temel performans göstergeleri tanımlanabilmelidir. Organizasyonel yapıya uygun şekilde temel performans göstergeleri arası ilişkiler kurulabilmeli, kişisel göstergelerden organizasyonel yapının üst katmanlarına doğru farklı ağırlıklarla toplanabilmelidir.
- Organizasyonel yapıyı yansıtacak şekilde “sorumluluk haritaları” tanımlanabilmeli, her göstergeye arzu edildiği takdirde farklı sorumlular atanabilmelidir.
- Belirlenen göstergelerin zaman içindeki gelişimini “trend grafikleri” ile grafiksel olarak gösterebilmek mümkün olmalıdır. Ayrıca, tanımlı göstergelerin gelecek dönemlerdeki gelişimi sistem tarafından tahmin edilebilmeli ve grafiksel ve tablo olarak Performans Yönetim sisteminde gösterilebilmelidir.
- Tanımlanan göstergeler ve ölçütler, karneler üzerinde biraraya getirilebilmelidir.
- Performans yönetim göstergeleri ve karneler, kurum yöneticileri için tek bir ekrandan sunulabilmeli ve “yönetici ekranı” tasarlanabilmelidir. Bu arayüz üzerinden, birden fazla karne, grafik ve analiz ekranı aynı ekranda gösterebilme imkânı olmalıdır.
- Tanımlanan gösterge veya karneler, herhangi bir ek programlamaya ihtiyaç kalmaksızın Microsoft Excel ve Microsoft PowerPoint uygulamalarına aktarılabilirdir.
- Performans yönetim sistemi bünyesinde, sorunlu veya sorunsuz görülen alanların detaylı olarak incelenebilmesi için analitik komponentler bulunmalıdır. Bu analitik komponentler ile farklı boyutlarda en üst seviyelerden en alt seviyeye kadar inen analizleri yapmak mümkün olmalı; neden-sonuç ilişkileri incelenebilmelidir.

Planlama ve Konsolidasyon:

- Planlama verileri Microsoft Excel üzerinden girilebilmelidir. Dolayısıyla sistem, Microsoft Excel üzerinden “write-back” işlemlerini desteklemelidir.

- Alınması istenen raporlar, Microsoft Excel ve web ortamından alınabilmelidir.
- Microsoft Excel üzerinde tüm departmanlara veri girişi veya raporlama amacıyla kullanılacak dosyalar, merkezi olarak saklanabilmeli ve yönetilebilimelidir.
- Son kullanıcılar, Microsoft Excel’de istedikleri Microsoft Excel formülü ile hazırlık yapabilmeli; istedikleri anda plan verilerini sunucuya gönderebilirler.
- Plan verileri girmek için kullanılan Microsoft Excel dosyaları, çevrimdışı çalışabilmelidir. Çevrimdışı çalışılan bir durumda, Microsoft Excel dosyası tekrar çevrimiçi olduğunda, verileri ve ilgili kuralları merkezi sunucu ile senkronize etme imkânı, planlama sistemi bünyesinde bulunmalıdır.
- Performans yönetimi sistemi, plan verilerini geçmiş verilerden yola çıkarak “tahmin edebilme” imkânına sahip olmalıdır. Bu özellik dahilinde, geçmiş dönemlerde oluşan verileri kullanıp istatistikî yöntemlerle gelecek veriler tahmin edilebilmeli, ve plan değeri olarak saklanabilmelidir.
- Bütçelemeye tanımlanabilecek modeller ve boyutlar açısından sistemde herhangi bir sınır bulunmamalıdır.
- Planlama sistemi, kurum bünyesinde var olan planlama ve bütçeleme iş akışlarını yansıtabilecek şekilde “iş akış” desteği içermelidir. İş akış desteği ile, yapılan bir planlamanın kabulü veya reddi mümkün olabilmeli, plan değerleri tanımlı kişiler tarafından gerektiği takdirde güncellenebilmelidir. Planlama iş akışları, kurum tarafından esnek şekilde ve herhangi bir programlama gerektirmeksizin tanımlanabilmelidir.
- Kurumun tamamını veya bir bölümünü ilgilendiren ortak iş kuralları ve plan değerleri, “varsayım” modeli şeklinde tanımlanabilmelidir. Varsayım modellerinde yer alan parametreler değiştirildiğinde kurumsal planların geneli ve/veya tamamı güncellenebilmelidir.
- Sistem, en iyi durum, en kötü durum, tahmini durum, v.b. farklı senaryoları oluşturmak için kullanılabilimelidir.
- Temel varsayım modelleri ile “enflasyon, şube sayısı, gelir vergisi oranı” gibi kurumsal parametreler tanımlanabilir. Bu parametreler (general assumptions), tüm kurumsal bütçelerin ortak bir zemine dayalı planlanmasını sağlar. Böylece, örneğin şube sayısı varsayımı değiştirildiğinde, tüm gelirler ve giderler, buna bağlı olarak da proforma finansal tablolar, belirli bir model çerçevesinde güncellenebilirler.

ETL Tool Seçimi:

- ETL yazılım aracı, veritabanından bağımsız bir engine'e sahip olmalı fakat gerektiğinde veritabanının sağlayacağı fonksiyonları da kullanabilmeli.
- ETL yazılım aracı, MS SQL Server, DB2, Oracle, Sybase, Sybase IQ gibi ilişkisel veritabanlarına native olarak erişebilmeli.
- ETL yazılım aracı sadece değişen kayıtların veri ambarına alınmasını sağlayacak mekanizmalara sahip olmalı (Changed Data Capture)
- Sadece değişen kayıtları veri ambarına alırken, ilişkisel veritabanının sağladığı olanaklar söz konusuysa, ilişkisel veritabanının loglarından bu bilgilere ulaşabilmeli.
- İlişkisel veritabanları dışında, Excel, Text, XML gibi farklı verikaynaklarından direk olarak veri çekebilmeli.
- XML verikaynaklarını okurken pipelining özelliği gösterebilmeli. Bu sayede büyük XML dosyaların tamamını bir defada okumak yerine parça parça okuyabilmek mümkün olmalı.
- Programlamaya gerek kalmadan sadece görsel geliştirme ortamında, ETL prosesleri oluşturulabilmeli.
- Yazılım aracının ürettiği kodlar gerektiğinde yetkisi olan kişiler tarafından elle müdahale edilebilecek özelliklere sahip olmalı.
- Üretilen kodlar, veritabanı seviyesinde çalışabileceği gibi gerektiğinde daha iyi performans sağlamak amacıyla ETL yazılım aracının sahip olduğu servisler üzerinde de çalışabilmeli.
- İlişkisel veritabanlarında varchar, long, cblob gibi field'lar içinde XML verinin tutulması durumunda bu XML sütunların ETL sürecine dahil edilebilmesi.
- Veri transformasyonu için hazır bileşenler ve built-in functions sunulmalı.
- Hazır bileşenler ve built-in function'lar veritabanına bir SQL cümlecığı gönderdiğinde hangi SQL cümleciklerinin gönderildiği yakalanabilmeli.
- Koşula bağlı transformasyon özelliği desteklenmeli.
- ETL programları geliştirilirken veritabanında direk çalıştırılması gereken fonksiyonlar belirlenebilmeli ve bu sırada ETL programı tarafından gönderilen değişkenler kullanılabilmesi.
- Veri ambarının modeli, ETL yazılım aracıyla kontrol edilebilmeli.
- Tabloların, indekslerin yaratılmasına izin veren bir altyapı mevcut olmalı.

- Veri ambarına ait metadata ilişkisel bir veritabanında saklanabilmeli.
- Veri ambarına ait metadata, sorgulama yazılım araçlarıyla erişilebilir nitelikte olmalı.
- Veri ambarlarında kullanılacak Star Schema ya da Snow flake gibi tasarımlar desteklenmeli.
- Oluşturulan kodlar ya da SQL cümlecikleri, daha sonra kullanılabilme amacıyla kaydedilebilmeli.
- Veri çekme sırasında sütun ya da satır bazında kısıtlama konulabilmeli.
- ETL yazılım aracı kendi kütüphanesinde bulunan transformasyonlara sahip olmalı.
- Oracle, DB2, Sybase, Sybase IQ, MS SQL Server gibi veritabanları için “bulk loader” özelliğini desteklemeli.
- Veri ambarında Real-time veri besleme için gerekli altyapıyı sunmalı.
- Diğer yazılım araçlarıyla entegrasyonu sağlayabilecek bir altyapıya sahip olmalı.
- Sorgulama raporlama araçlarıyla entegre çalışabilmeli.
- Kaynak ve hedef veritabanlarındaki fiziksel verileri izleyebilecek ve örnek verileri görüntüleyebilecek özellikleri olmalı.
- Transformasyonlar sırasında veritabanından alınan veriyi daha sonraki adımlarda kısıtlayabilmeli ve gerektiğinde sıralayabilmeli.
- ETL scriptlerinin yeniden kullanılabilmesine (Object Reusability) izin veren bir altyapı olmalı ve gerektiğinde önceden hazırlanan kodlar tek bir bileşen gibi tekrar tekrar kullanılabilmeli.
- Herhangi bir tablo ETL scriptleri içinde hem kaynak(source) hem hedef(target) olarak kullanılabilmeli.
- Hedef tabloya yüklenecek veri kümeleri üzerinde validasyon yapmayı sağlayacak hazır fonksiyon ve/veya transformasyon bileşenleri olmalı.
- Geliştirme ortamında profiling özelliği olmalı. Örneğin, ETL yazılımcısı, bir tabloya ait bir sütunda kaç farklı değer olduğu, minimum ve maximum değerlerinin ne olduğu, % kaçlık bir bilginin boş ya da NULL olduğu, ya da bu sütundaki değerlerin tüm tablo içinde % kaçlık bir kısıtlı olduğunu görebilmeli.
- Relationship profiling desteği olmalı. Örneğin aralarında ilişki olan 2 tablo seçildiğinde, 1. tabloda olan 2. tabloda olmayan ya da tam tersi 2. tabloda olan ve 1. tablo da olmayan ve her ikisinde birden olan kayıtların neler olduğu gerek yüzdesel olarak gerek kayıt bazında izlenebilmeli.
- Metadata yönetimine izin vermeli.

- Veri ambarında bir sütun değiştirildiğinde bundan hangi kurumsal raporların etkilenebileceğini gösterecek analizleri sunabilmeli.
- ETL prosesleri için gruplar oluşturarak daha sonra schedule edebilmeli.
- ETL yazılım aracının Scheduler'ı kullanılabileceği gibi, daha farklı bir scheduler'ın kullanılması durumunda (Unix Cron ya da Windows Scheduler gibi) ETL scriptleri de bu scheduler'a gönderilebilmeli.
- Lokal ve global değişkenler tanımlanabilmeli.
- Metadata bilgileri raporlanabilmeli.
- Bir proses başarısızlıkla sonuçlandığında otomatik düzeltmeleri yapabilmeli (automatic error&recovery).
- ETL yazılım aracı kendi motoruna sahip olmalı, gerektiğinde ayrı veritabanının olduğu sunucu dışında herhangi bir sunucuya kurulabilmelidir. Bu sayede veritabanına yük getirmeden çalıştırılacak bir ortam sağlanmalıdır.
- ETL yazılım aracı gerektiğinde ELT fonksiyonlitesini sağlayacak şekilde kurulabilmelidir.
- Bir proses başarısızlıkla sonuçlandığında elle düzeltmelere izin vermeli (manual recovery)
- Başarısızlıkla sonuçlanan bir ETL prosesinde, başarısızlığa neden olan verileri bir tabloya atabilmeli.
- Başarısızlıkla sonuçlanan bir ETL prosesinde, başarısızlığa neden olan SQL cümleciklerini kaydedebilmeli.
- Çok kullanıcıya geliştirmeye izin veren bir altyapıya sahip olmalı.
- ETL proseslerini izlemeye izin vermeli.
- Başarısızlıkla sonuçlanan bir iş olduğu zaman ilgili kişilere otomatik email gönderebilmeli.
- Paralel çalıştırma özelliklerini hem kaynak hem de hedef tablolar üzerinde destekleyebilmeli.
- Transformasyonların işlenmesi sırasında paralellik sağlanabilmeli (transformasyonlar için birden fazla thread yaratma gibi...).
- Text dosyaları okurken de paralel yükleme özelliklerini destekleme.
- Text dosyalar direk file sistemden okunabileceği gibi FTP gibi bir program aracılığıyla da erişilerek kaynak olarak kullanılabilmeli ya da gerektiğinde output olarak yine bir FTP site'ına yüklenebilmelidir. Bu işlem schedule edilebilmeli.

- Belli bir directory’de birden fazla Text dosyanın bulunması durumunda, ETL yazılım aracına bu dosyaların isimleri teker teker verilmeden sadece directory adı verilerek bu directory’deki tüm dosyaların alınması ve işlenmesi sağlanabilmelidir.
- Başarısızlıkla sonuçlanan bir ETL prosesi olduğunda, bu proses daha sonra yeniden çalıştırıldığında aynı veriyi yeniden yüklememeli.
- While döngülerini desteklemeli.
- Exception Handling mekanizması olmalı.
- Bir ETL prosesi veriden dolayı başarısızlıkla sonuçlandığında, bu veri için otomatik overflow dosyalar oluşturabilmeli.
- Bir boyuta sahip olmayan hesaplamaları yükleyebilmek için gerekli mekanizmalara sahip olmalı.
- Metadatayı diğer yazılım araçlarıyla paylaşabilecek altyapıya sahip olmalı.
- Test, geliştirme, üretim gibi farklı ortamları desteklemeli.
- Fiziksel sütun adları yerine mantıksal sütun adları verilebilmeli.
- İstatistiki log bilgileri sunabilmeli.
- Hata log bilgileri sunabilmeli.
- ETL proseslerini izleyebilecek log bilgilerini sunabilmeli.
- Geliştirilen ETL scriptlerinin farklı versiyonlarının birbiriyle karşılaştırılması ve aradaki farkların listelenmesi mümkün olmalı.
- farklı tablolar arasındaki yapısal değişiklikler listelenebilmeli (Tablo1 ve Tablo2’nin sahip olduğu field’lar ve bunların tiplerinin aynı olup olmadığının karşılaştırılması).
- Built-in fonksiyonlar sunmalı.
- Stored procedure’ların kullanımına izin vermeli.
- BeforeTableLoad, AfterTableLoad gibi bir takım event’leri desteklemeli.
- Kaynak ve hedef tablolar arasında karşılaştırmaya izin veren transformasyon bileşenleri olmalı.
- ETL prosesleri oluşturulurken yapılanlara ait not yazılabilecek açıklama hücreleri oluşturulabilmeli.
- Sistemde ETL scriptlerini debug edebilecek altyapı bulunmalı.
- Verileri tanımak için profiling yapabilecek, bir sütundaki farklı bilgilerin neler olduğunu bunların yüzdeleriyle birlikte verebilecek özellikler olmalı.
- ERWIN/Power Designer gibi modelleme araçlarıyla entegre olabilmeli.

5.10. İş Zekası Raporlama ve Dashboard Örnekleri

Aşağıda bankacılık sektörü için İş Zekası platformu kullanarak hazırlanmış olduğum raporlama ve dashboard örnekleri gösterilmiştir.

Aşağıda Şekil 5.10. da ve Şekil 5.11. de bankacılık sektörüne ait Statik Raporlama Örnekleri gösterilmiştir.

Belirli bir tarih itibarıyla açılan Vadeli hesaplar Raporu

Hesap No	Adı	HT	MT	Döviz	Valör	Vade	Toplam Bakiye	Faiz	Açıklama
902700002468370	TOROS TARIM SAN.VE T	SK	GD	EUR	26/05/2009	27/05/2009	217,000.00	3.00	Temdit Edilmeyecek
902700026122544	EFG FİN.KİRALAMA A.Ş	SK	GD	EUR	26/05/2009	27/05/2009	1,343,000.00	3.00	Temdit Edilmeyecek
902700142977367	HAYDAR AGCET	NU	GA	EUR	26/05/2009	30/06/2009	2,690.00	3.25	Temdit Edilmeyecek
902700220762455	VALUE PARTNERS	SK	GD	EUR	26/05/2009	27/05/2009	160,035.00	2.50	Temdit Edilmeyecek
902700224979515	HASAN BULUT	SC	GA	EUR	26/05/2009	27/05/2009	550,000.00	3.50	Temdit Edilmeyecek
904300144774360	HİMMET KOLCU	NU	GA	EUR	26/05/2009	30/06/2009	548.77	3.00	Anapara+Faiz
904300190943354	LÜTFİYE KARDEMİR	NU	GA	EUR	26/05/2009	29/06/2009	7,215.82	3.25	Anapara+Faiz
906800027428595	MAĞDENLİ A.Ş	NU	GD	EUR	26/05/2009	29/06/2009	200,000.00	4.00	Anapara+Faiz
909200116544393	MUAMMER ÖCAL	NU	GA	EUR	26/05/2009	29/06/2009	10,492.00	3.25	Temdit Edilmeyecek
909200116662357	NUR LATİFOĞLU	NU	GA	EUR	26/05/2009	29/06/2009	729.00	3.25	Temdit Edilmeyecek
909200172955380	AYŞE GÜREL	NU	GA	EUR	26/05/2009	29/06/2009	2,344.00	3.25	Temdit Edilmeyecek
909200196750418	NUR MAÇ	NU	GA	EUR	26/05/2009	29/06/2009	1,052.00	3.25	Temdit Edilmeyecek
909200209535369	ERDAL AŞÇILI	NU	GA	EUR	26/05/2009	29/06/2009	1,285.00	3.25	Temdit Edilmeyecek
910000144020462	YILMAZ KUZUCU	NU	GA	EUR	26/05/2009	29/06/2009	100,000.00	4.00	Temdit Edilmeyecek
910000144742554	SADIYE AŞİCİ	NU	GA	EUR	26/05/2009	29/06/2009	3,550.00	3.25	Temdit Edilmeyecek
910000159080402	ŞAHİSMAIL GÖKÇE	NU	GA	EUR	26/05/2009	29/06/2009	2,662.00	3.25	Temdit Edilmeyecek
911800187797371	MERAL DEMİR TÜRK	NU	GA	EUR	26/05/2009	26/08/2009	6,435.65	3.75	Anapara+Faiz
912600010694611	HÜRMÜZ AYSU ARIBURNU	NU	GA	EUR	26/05/2009	30/06/2009	9,145.00	3.25	Temdit Edilmeyecek
919100138954402	ATİYE ÇİĞDEM ACAR	NU	GA	EUR	26/05/2009	29/06/2009	4,410.00	3.25	Temdit Edilmeyecek
921300155524413	SEVİM HAFİZE KÖROĞLU	NU	GA	EUR	26/05/2009	29/06/2009	51,000.00	3.25	Temdit Edilmeyecek
921300160472557	AYÇA ERTUĞ	NU	GA	EUR	26/05/2009	29/06/2009	898.00	3.25	Temdit Edilmeyecek
921300184479436	MEHMET GÜNNASIR	NU	GA	EUR	26/05/2009	29/06/2009	2,347.00	3.25	Temdit Edilmeyecek
921300191748399	HASAN KİRAZ	SC	GA	EUR	26/05/2009	02/06/2009	4,090.00	1.50	Anapara+Faiz
922400166065480	KAZIM ÇALIŞKAN	NU	GA	EUR	26/05/2009	29/06/2009	2,356.70	3.25	Temdit Edilmeyecek
922400212152389	RABIYE ÇAĞLAYAN	SC	GA	EUR	26/05/2009	02/06/2009		3.50	Temdit Edilmeyecek

Records 1 - 25

Şekil 5.10. Statik Raporlama Örnek-1

Vadesi Gelecek Vadeli Hesaplar Raporu

Hesap No	Adı	Ürün	HT	MT	Döviz	Valör	Vade	Toplam Bakiye	Faiz	Açıklama
903500142540373	NEDİM DEMİRHAN	Vadeli Mevduat TL	LK	BA	TRY	20/05/2009	27/05/2009	15,074.92	9.75	Anapara+Faiz
					TRY Toplam			15,074.92		
902700002468368	TOROS TARIM SAN.VE T	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	21,048,400.00	10.25	Temdit Edilmeyecek
902700005684365	ERMETAL A.Ş.	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	250,000.00	9.00	Temdit Edilmeyecek
902700026122542	EFG FİN.KIRALAMA A.Ş.	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	250,000.00	9.25	Temdit Edilmeyecek
902700142432589	TEKFEN HOLDİNG A.Ş.	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	990,000.00	9.25	Temdit Edilmeyecek
902700143693397	TEKFEN ENDÜSTRİ VE T	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	110,000.00	9.25	Temdit Edilmeyecek
902700144249453	TEKFEN SİGORTA ARACI	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	952,000.00	9.25	Temdit Edilmeyecek
902700145313429	TEKFEN TURİZM VE İŞL	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	660,000.00	9.25	Temdit Edilmeyecek
902700165159455	TEKFEN EMLAK GEL. YA	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	325,000.00	9.25	Temdit Edilmeyecek
902700193894396	BAŞTUĞ A.Ş.	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	389,000.00	9.25	Temdit Edilmeyecek
902700198343520	EFG İSTANBUL MENKUL	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	2,970,000.00	9.25	Temdit Edilmeyecek
902700211351426	TEKFEN OZ GAYRİMENKU	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	1,248,000.00	9.25	Temdit Edilmeyecek
906800083593415	SAFİR LTD ŞTİ	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	107,000.00	8.75	Temdit Edilmeyecek
908400227951367	SEÇKİN ONUR A.Ş.	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	1,829,000.00	10.50	Temdit Edilmeyecek
910000226891389	BİA DIŞ TİC. A.Ş.	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	659,000.00	8.75	Temdit Edilmeyecek
911800060394380	BUDİNOKS LTD ŞTİ	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	250,305.77	8.75	Anapara+Faiz
911800226586358	ÖZEL MAYA SAĞ.HİZ Tİ	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	400,633.95	9.00	Anapara+Faiz
925500182109363	ORBIT MÜH.İNŞ.SATIN.	Vadeli Mevduat TL	LK	BB	TRY	20/05/2009	27/05/2009	25,000.00	9.75	Temdit Edilmeyecek
927800172764477	PLASTİKKART AKİLLİKA	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	300,985.15	9.75	Anapara+Faiz
934000186933357	KUMSAN ASFALT LTD_ŞT	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	301,759.51	9.50	Anapara+Faiz
937000147186358	İNTRA A.Ş.	Vadeli Mevduat TL	LK	BB	TRY	25/05/2009	27/05/2009	160,000.00	8.50	Temdit Edilmeyecek
941000226134403	MERGRAIN LTD.ŞTİ.	Vadeli Mevduat TL	LK	BB	TRY	20/05/2009	27/05/2009	300,000.00	12.15	Temdit Edilmeyecek
941000226134404	MERGRAIN LTD.ŞTİ.	Vadeli Mevduat TL	LK	BB	TRY	20/05/2009	27/05/2009	300,000.00	12.15	Temdit Edilmeyecek
941000226134405	MERGRAIN LTD.ŞTİ.	Vadeli Mevduat TL	LK	BB	TRY	20/05/2009	27/05/2009	238,242.49	12.15	Temdit Edilmeyecek
943000208224587	MEGA METALÜRJİ DEM.S	Vadeli Mevduat TL	LK	BB	TRY	26/05/2009	27/05/2009	200,000.00	9.25	Temdit Edilmeyecek

Records 1 - 25

Şekil 5.11. Statik Raporlama Örnek-2

Aşağıda Şekiller 5.12. / 5.13. ve 5.14 da bankacılık sektörüne ait Drill-Down özelliği olan bir raporlama örneği gösterilmiştir.

Drill - Down özelliği olan raporlama örneği (Üç detayda – Segment/Şube/Müşteri)

Segment Bazında Menkul Kıymet Emanet Stoğu
* Lütfen Segment Detayını Görebilmek İçin Segment Adına Tıklayınız!

Segment Kodu	Ürün	Nominal	Tutar TL
Genel Toplam		187,545,803.00	149,234,001.82
LC	KESİN	1,321,500.00	1,055,253.45
	REPO	56,695.00	54,300.00
LC Toplam		1,378,195.00	1,109,553.45
RTL	KESİN	168,364,101.00	133,692,034.72
	REPO	1,306,119.00	1,250,943.66
RTL Toplam		169,670,220.00	134,942,978.38
SBB	KESİN	3,029,533.00	2,516,637.34
	REPO	432,421.00	414,150.00
SBB Toplam		3,461,954.00	2,930,787.34
SME	KESİN	11,638,070.00	8,912,351.74
	REPO	1,397,364.00	1,338,330.91
SME Toplam		13,035,434.00	10,250,682.65

Şekil 5.12. Drill - Down Raporlama Örneği - Segment

LC Segmentine tıklandığında açılan detay rapor aşağıdadır (LC segmentine ait şubeleri içeren rapor).

Şube Bazında Menkul Kıymet Emanet Stoğu
* Şube Detaylarını Görebilmek İçin Şube Koduna Tıklayınız!

Segment Kodu	Şube Adı	Kodu	Ürün	Nominal	Tutar TL
Genel Toplam				1,378,195.00	1,109,553.45
LC	ANTALYA	9213	REPO	52,205.00	50,000.00
	ETİLER	9191	KESİN	132,900.00	157,171.56
	MERKEZ	9027	REPO	4,490.00	4,300.00
	TAKSİM	9370	KESİN	1,188,600.00	898,081.89
LC Toplam				1,378,195.00	1,109,553.45

[Return](#) - [Download](#) - [Create Bookmark Link](#)

Şekil 5.13. Drill - Down Raporlama Örneği - Şube Detayı

Taksim şube koduna tıklandığında o şube altındaki müşterileri gösteren rapor örneği aşağıdadır.

 **Müşteri Bazında Menkul Kıymet Emanet Stoğu**

Segment Kodu	Şube Adı	Kodu	Müşteri Adı	Ürün	Nominal	Tutar TL
Genel Toplam					1,188,600.00	898,081.89
LC	TAKSİM	9370	YAPI KREDİ EMEKLİLİK	KESİN	1,188,600.00	898,081.89
LC Toplam					1,188,600.00	898,081.89

Şekil 5.14. Drill-Down Raporlama Örneği - Müşteri Detayı

Aşağıda Şekiller 5.15. / 5.16. ve 5.17 de bankacılık sektörüne ait Analitik Raporlama örnekleri gösterilmiştir.

(26/05/2009 / 25/05/2009)

Cari Mevduat TL

Günlük Bakiye	Önceki Gün Bakiye	Günlük Artış / Azalış	Günlük Artış / Azalış (%)	Önceki Ay Ortalama	Günlük Ortalama
8,889,959.53	9,735,109.05	-885,149.52	-8.89	8,817,028.30	9,133,349.18

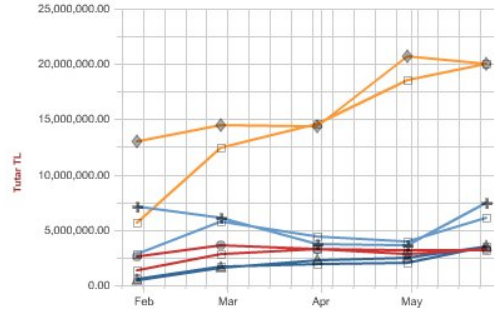
(26/05/2009 / 25/05/2009)

Cari Mevduat TL

Şube Adı	Şube Kodu	Günlük Bakiye	Önceki Gün Bakiye	Günlük Artış / Azalış	Günlük Artış / Azalış (%)	Önceki Ay Ortalama	Günlük Ortalama
ADANA	9297	130,084.89	241,822.39	-111,737.70	-46.21%	191,097.40	176,887.59
ALANYA	9380	279,301.95	279,818.05	-516.10	-0.18%	128,139.34	175,906.88
ALTUNIZADE	9255	483,234.72	443,895.25	39,339.47	8.86%	285,019.21	465,713.87
ANKARA	9092	98,341.78	77,971.19	20,370.59	26.13%	94,930.89	98,627.52
ANTALYA	9213	149,059.53	146,993.64	2,065.89	1.41%	135,996.52	147,784.50

Şekil 5.15. Analitik Raporlama Örnek-1

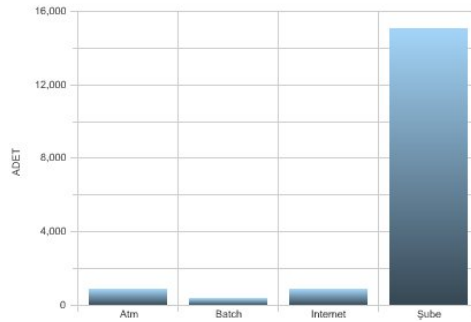
Segment Kodu	Ürün	30/01/2009		27/02/2009		31/03/2009		30/04/2009		26/05/2009	
		Ortalama	Nokta	Ortalama	Nokta	Ortalama	Nokta	Ortalama	Nokta	Ortalama	Nokta
LC	EFG A DEGISKEN			0,00	0,00	0,00	0,00	1,000.15	1,000.15	1,050.48	1,087.04
	EFG A HISSE			0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	EURO DEGIŞ	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	EURO LIKİT	511,639.42	502,689.23	1,685,495.58	1,617,073.54	1,934,361.38	2,277,814.73	1,986,095.83	2,470,856.33	3,517,448.72	3,481,765.46
LC Toplam		511,639.42	502,689.23	1,685,495.58	1,617,073.54	1,934,361.38	2,277,814.73	1,986,095.83	2,470,856.33	3,517,448.72	3,481,765.46
RTL	EFG A DEGISKEN			4,730.55	12,028.81	34,588.51	249,874.15	449,087.41	335,394.17	583,431.38	919,584.23
	EFG A HISSE			22,571.54	73,042.20	82,233.69	92,019.94	334,118.09	324,012.85	284,202.48	281,880.49
	EURO DEGIŞ	229,752.17	604,298.58	1,134,689.88	1,374,074.85	1,388,228.51	1,571,180.58	2,137,485.95	2,295,098.80	2,594,847.13	2,304,359.13
	EURO LIKİT	5,389,133.16	12,373,086.12	11,248,928.75	13,024,295.19	13,132,479.57	12,424,838.59	15,590,420.88	17,754,162.97	16,529,840.88	16,487,599.07
RTL Toplam		5,618,895.33	12,977,384.70	12,410,920.53	14,483,441.05	14,637,525.28	14,337,513.26	18,511,092.13	20,708,636.59	19,972,121.65	19,973,822.92
SBB	EFG A DEGISKEN			0,00	0,00	0,00	0,00	557.11	551.59	2,059.83	18,802.52
	EFG A HISSE			0,00	0,00	0,00	0,00	6,103.20	6,408.39	16,860.18	38,687.96
	EURO DEGIŞ	1,584.27	3,047.89	36,561.95	88,578.46	71,347.81	46,110.95	48,813.35	49,721.71	38,416.05	5,894.21
	EURO LIKİT	1,311,732.84	2,549,555.02	2,809,230.71	3,478,908.29	3,236,345.58	3,251,937.16	3,094,012.44	2,782,198.27	3,092,008.44	3,183,892.88
SBB Toplam		1,313,297.11	2,552,602.91	2,845,792.66	3,567,486.75	3,307,693.37	3,298,048.11	3,149,486.10	2,838,889.96	3,119,345.30	3,247,077.37
SME	EFG A DEGISKEN			0,00	0,00	0,00	0,00	1,000.15	1,000.15	2,827.91	3,183.15
	EFG A HISSE			0,00	0,00	0,00	0,00	2,039.75	2,079.20	5,174.23	17,568.64
	EURO DEGIŞ	2,981.63	5,809.51	23,958.84	32,651.99	49,890.64	33,121.59	60,291.19	83,881.85	2,544,375.77	2,566,487.72
	EURO LIKİT	2,822,283.17	7,159,817.84	5,706,047.93	6,081,777.05	4,400,650.74	3,681,109.74	3,841,718.18	3,507,684.89	3,540,988.05	4,907,889.91
SME Toplam		2,825,264.80	7,165,627.35	5,730,004.77	6,114,429.04	4,450,441.38	3,714,231.33	3,905,049.27	3,594,645.89	6,093,363.96	7,495,109.42



Şekil 5.16. Analitik Raporlama Örnek-2

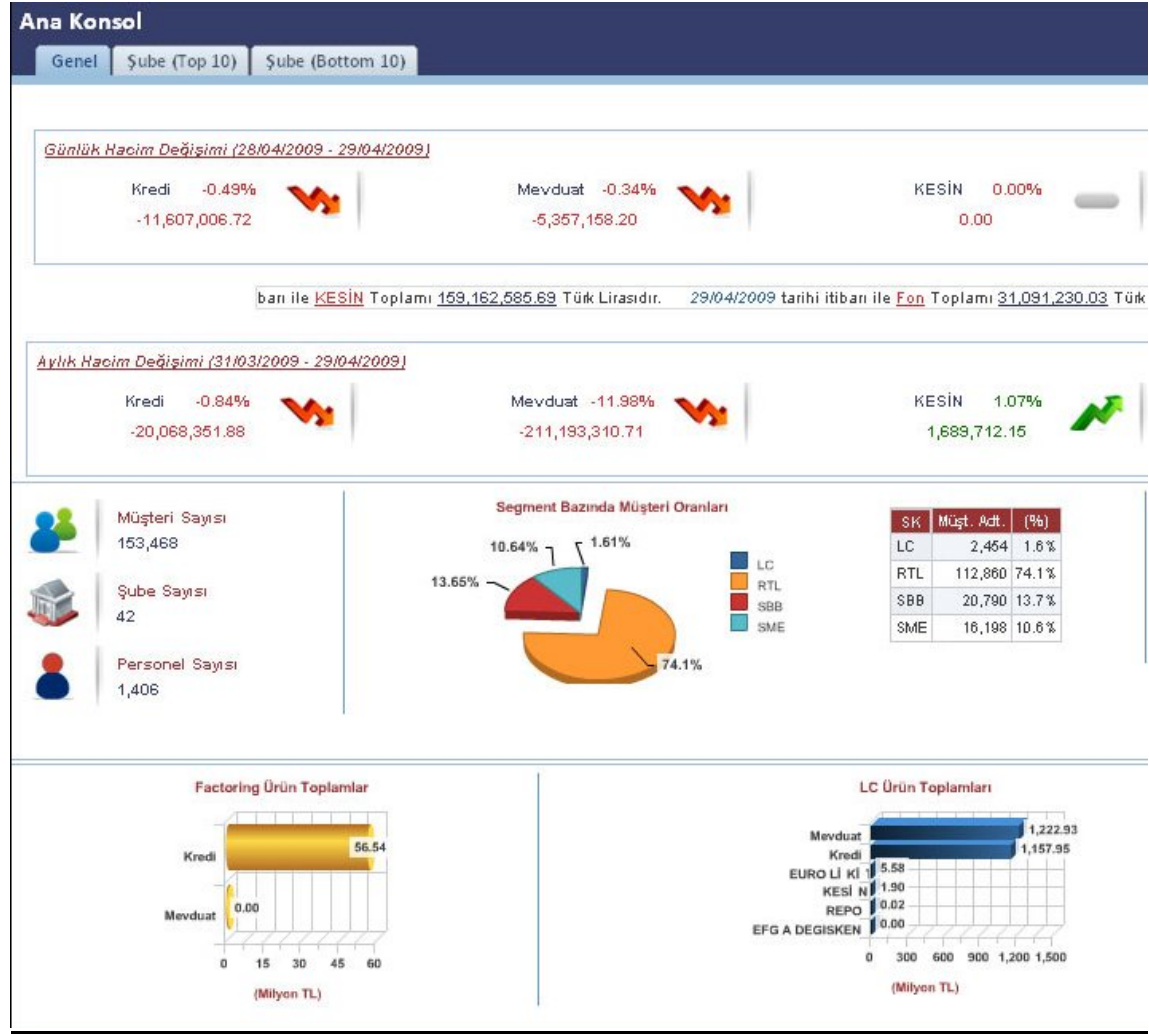
Tarih Seçiniz! Between and Şube Seçiniz! Kasa No Seçiniz! Modül Seçiniz! İşlem Seçiniz!

KANAL ADI	ADET	HACİM (TL)	HACİM (EUR)
Atm	854	236,602.04	109,720.78
Batch	385	542,466,156.63	251,561,007.49
İnternet	860	4,521,498.47	2,096,780.32
Şube	15,057	4,576,177,927.42	2,122,137,794.99
Toplam	17,156	5,123,402,184.56	2,375,905,303.58

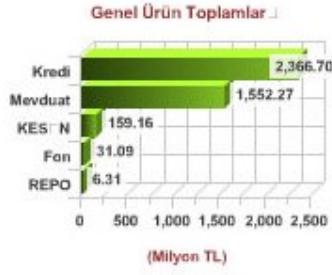


Şekil 5.17. Analitik Raporlama Örnek-3

Aşağıda bankacılık sektörüne ait Dashboard örnekleri gösterilmiştir.



Şekil 5.18. Dashboard Örnek-1



NEW Kur Bilgileri Konsolu Sisteme eklenmiştir!

Konsola ulaşmak için [Tıklayınız!](#)

NEW Aylık ve Haftalık Ortalama Raporları Sisteme eklenmiştir!

Raporlara ulaşmak için [Tıklayınız!](#)



Şekil 5.19. Dashboard Örnek-2

Musteri Bilgileri Analizi

Segment Bazında

Bölgesel Bazda

Kanal Bazında

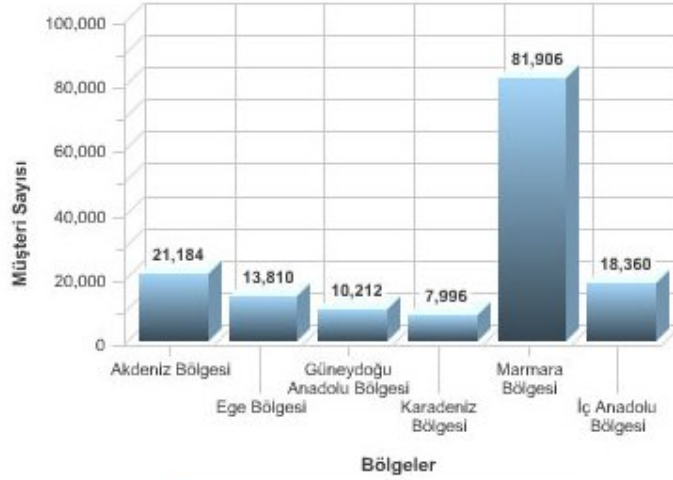
Tanım Tarihleri Bazında

Faaliye



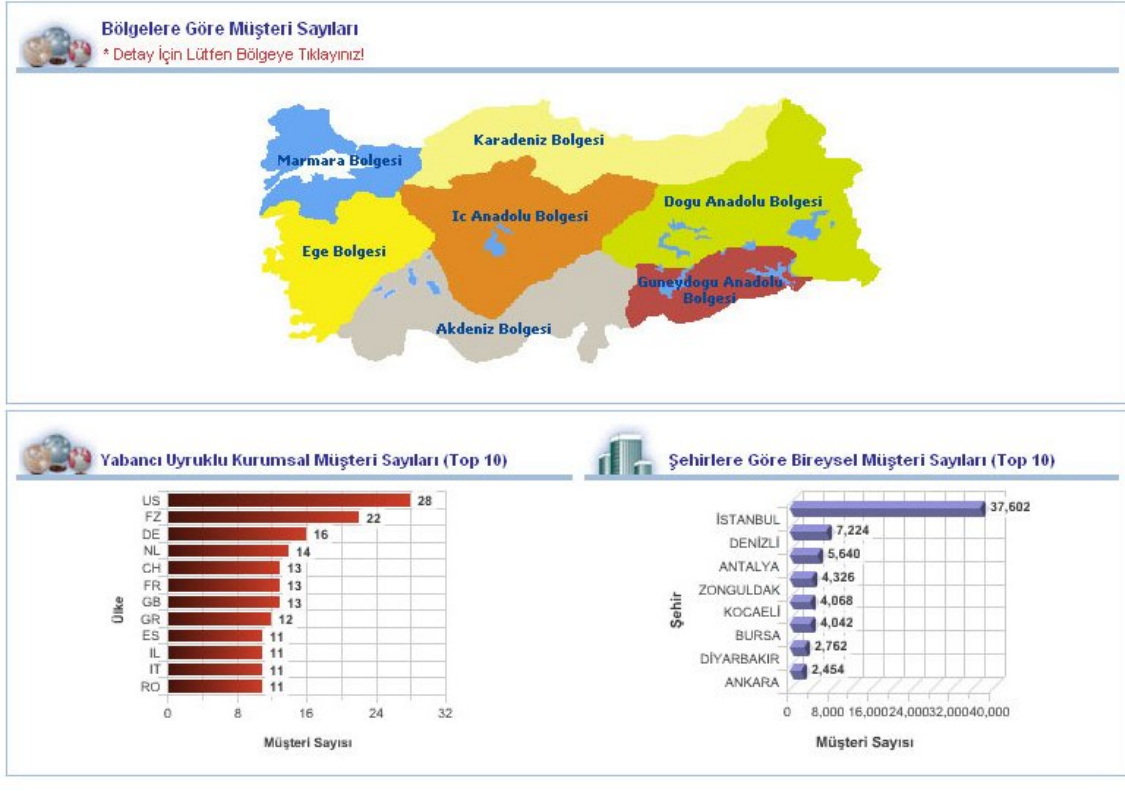
Bölgesel Bazda Müşteri Sayıları

* Detay Görmek İçin Grafik veya Tablo Üzerindeki Bölgeye Tıklayınız!



Bölgeler	Müşteri Sayısı
Akdeniz Bölgesi	21,184
Ege Bölgesi	13,810
Güneydoğu Anadolu Bölgesi	10,212
Karadeniz Bölgesi	7,996
Marmara Bölgesi	81,906
İç Anadolu Bölgesi	18,360
Genel Toplam	153,468

Şekil 5.20. Dashboard Örnek-3



Şekil 5.21. Dashboard Örnek-4

SONUÇLAR

Çalışmada kurumsal iş zekası tüm katmanları ile birlikte incelenmeye çalışılmış, teorik bilgilendirme ve uygulama örnekleri ile kurumlar açısından iş zekasının sağladıkları, kazanımları, amacı anlatılmıştır. Çalışmada iş zekası tanımları, kullanım amaçları, raporlama tanımları, sağladıkları, kazanım ve gelişim bilgileri ifade edilmiştir. Veri Ambarı ve kurulumu tüm detayları ile incelenmiş ve bu çalışmada teorik ve örneklemelerle bilgi verilmeye çalışılmıştır.

Veri ambarcılığı konusunda gözden kaçırılmaması gereken ve belki de en önemli nokta, oluşturulacak bir veri ambarının, ait olduğu kurumun özelliklerine ve gereksinimlerine bağlı olacağı, dolayısıyla herhangi bir standart yapıya oturtulmasının çok güç olduğudur.

Veri ambarı oluşturma çalışmalarının başarısını belirleyecek en önemli etkenlerden birisi, veri ambarının modellenmesidir. Bu çalışmada veri modelleme süreci ve veri ambarının yaşam döngüsü için çok önemli olan veri modelinin oluşturulması konusundaki yöntemler hakkında teorik ve pratik bilgiler verilmeye çalışılmıştır. Bu çalışmada veri ambarı oluşturulma süreçlerinden oldukça detaylı bir şekilde bahsedilmiştir. Bankacılık sektöründeki uygulamaları ile örneklendirilmiştir.

Çalışmada veri madenciliğinin tanımı, işlevi, yararları, bilgi keşfi, yöntemleri incelenmiş ve kurumlar açısından getirileri ve kullanım alanları anlatılmıştır. Bankacılık sektöründeki kullanım alanları, yöntemleri ve özellikleri örneklendirilmiştir.

Ayrıca bu çalışmada kurumsal performans yönetimi ve kurumsal karneleme konuları üzerine bilgilendirme yapılmıştır.

İş zekası platform ETL tool seçiminde dikkat edilmesi gereken konular ve soru setleri bu çalışmanın içeriğine koyulmuştur.

KAYNAKÇA

Almuallim, H. and Dietterich, T., 1991, Learning with many irrelevant features, In Proceedings of AAI 91, (Menlo Park, CA), AAAI Press, pp. 547-552.

Altun, H., 2001. ‘‘Comptabilité à Base d’Activité’’; Yüksek Lisans Tezi; Galatasaray Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.

Bain, T., Benkovich, M., Dewson, R., Ferguson, S., Graves, C., Joubert, J.T., Lee, D., Scott, M., Skoglund, R., Turley, P., Youness, S., 2000. Professional SQL Server 2000 Data Warehousing with Analysis Services, (Wrox Pres).

Ballard, C., Herreman, D., Schau, D., Bell, R., Kim, E. and Valencic, A., 1998, August, February, Data modeling techniques for data warehousing, System Services Corporation, Chicago.

Deogun, J. S., Raghavan, V. V. and Sever, H., 1995, August, Exploiting upper approximations in the rough set methodology, In The First International Conference on Knowledge Discovery and Data Mining (U. Fayyad and R. Uthurusamy, eds.), (Montreal, Quebec, Canada), pp. 69-74.

Fayyad, P. S. U. M., Piatetsky-Shapiro, G. and Uthurusamy, R., 1996a, Advances in knowledge discovery and data mining, Cambridge, MA: MIT Press.

Fayyad, P.S.U. M., Piatetsky-Shapiro, G., 1996b, The KDD process for extracting useful knowledge from volumes of data, Communications Of ACM 39, 11, 27-34.

Gupta, V. R., 1997, August, An introduction to data warehousing, System Services Corporation, Chicago.

Kimball, R., Reeves, L., Ross, M., Thornthwaite, W., 1998. The Data Warehouse Lifecycle Toolkit, (John Wiley & Sons, Inc.).

Kira, K. and Rendeli, L., 1992, The feature selection problem; Traditional methods and a new algorithm, In Proceedings of AAAI 92, AAAI Press pp. 129-134.

Lane, P., Schupmann, V., 2002. Oracle9i Data Warehousing Guide, Oracle Pres.

Pawlak, Z., Slowinski, K., and Slowinski, R., 1986, Rough classification of patients after highly selective vagotomy for duodenal ulcer, International Journal Of Man-Machine Studies, vol. 24, pp. 413-433.

Quinlan, J. R., 1986, Induction of decision trees, Machine Learning, vol. 1, pp. 81-106.

Oracle 10g : Data Warehousing Fundamentals, 2006 August.

Yarımağan, Ü., 2000, Veri tabanı sistemleri, Akademi & Türkiye Bilişim Vakfı, Ankara.

İnternet Kaynakları

Alpaydın, E., 1999, Zeki veri madenciliği: Ham veriden altın bilgiye ulaşma yöntemleri, www.cmpe.boun.edu.tr/~ethem/files/papers/veri-maden_2k-notlar.doc. (20.08.2008 - 20.11.2008)

Boar, B., 2000, December 25, Understanding data warehousing strategically, www.carleton.com.au/Understanding%20Data%20Warehousing%20Strategically.htm (05.08.2008 – 10.11.2008)

ÖZGEÇMİŞ

Malatya ili Hekimhan ilçesi doğumluyum. İlkokulu Nazilli Recepbey İlkokulunda, ortaokulu Nazilli Atatürk Ortaokulu ve liseyi ise Nazilli Lisesi'nde tamamladıktan sonra, Lisans öğrenimimi Anadolu Üniversitesi, Fen-Edebiyat Fakültesi, Matematik Bölümü'nde 1986 yılında tamamladım. Daha sonra 2006 yılında, Beykent Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı'nda yüksek lisans eğitimine başladım.

1988 yılından beri bankaların IT departmanlarında çeşitli kademelerde görev yaptım. Halen bir yabancı bankada IT departmanında CRM, İş Zekası, Bütçe ve Planlama, MIS konularından sorumlu yöneticiyim.

Yabancı dilim İngilizce olup, evli ve bir çocuk babasıyım.

Aday : Yunus Nadi ERDEMİR