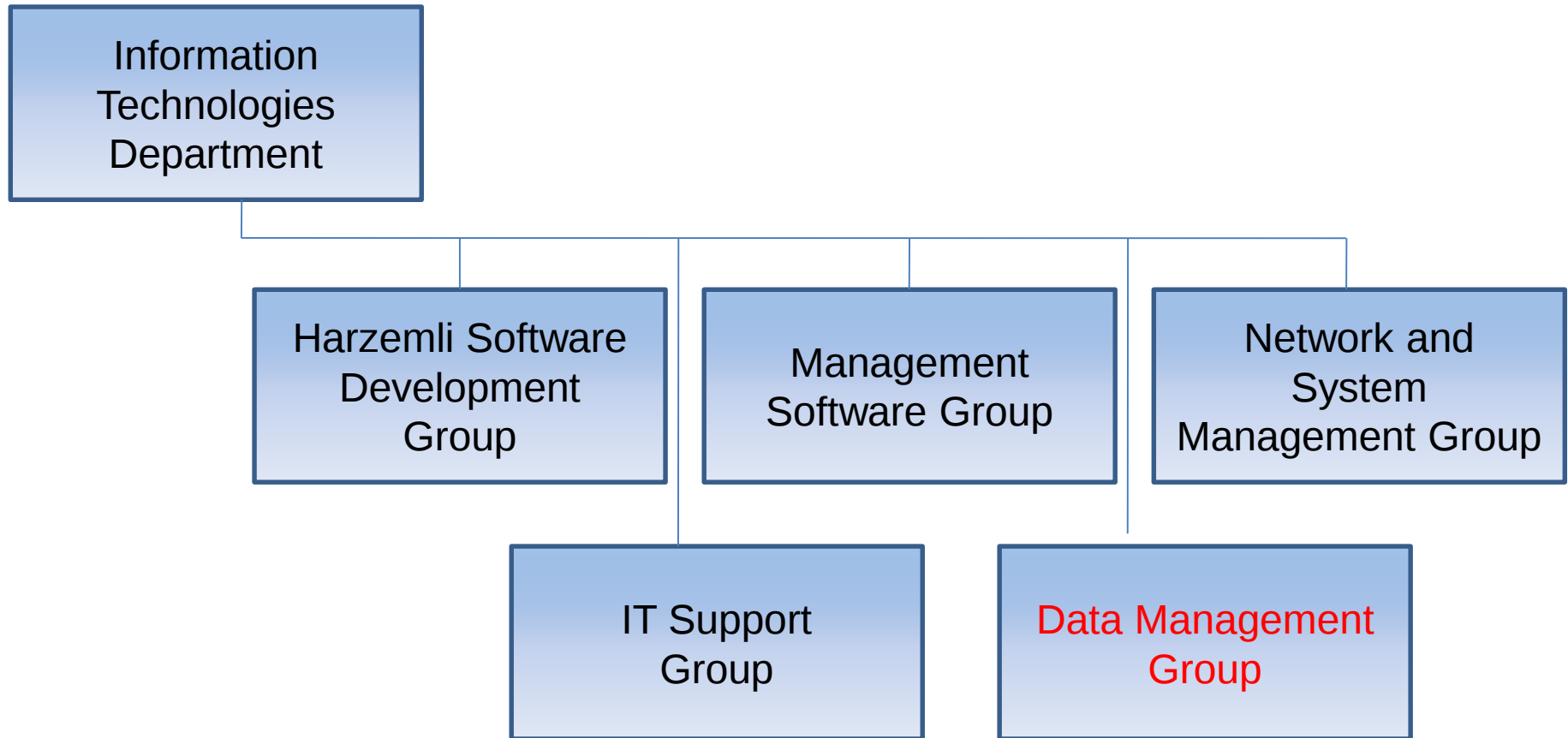


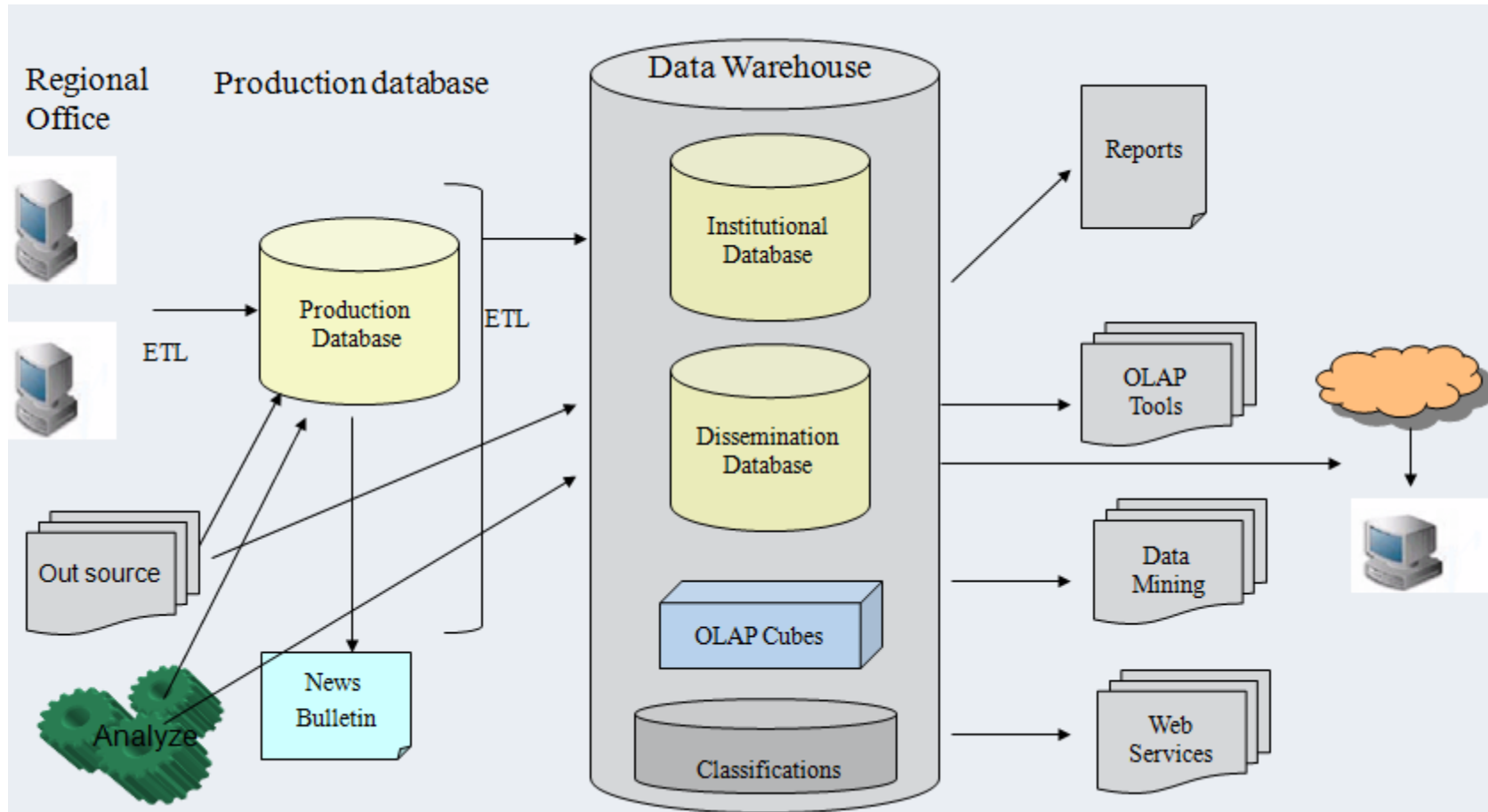
DATA PROCESS and Data Warehouse Projects in TURKSTAT

***31.03.2014
(Muscat, Oman)***

Information Technologies Department



TurkStat Data Architecture



TurkStat Databases

Production Database :

Current data is stored and always transaction in it.

Institutional Database (Enterprise Database):

The main repository of TurkStat's harmonized historical data.

Dissemination Database :

Aggregated time series data created from TurkStat institutional databases.

Up to now there are

62 institutional databases

59 dissemination databases

Oracle DBMS is used for all database

Production Databases

- Mainly data entry systems
- Database for web applications
- OLTP(Online Transactional Processing)
- Surveys data are stored in this database
- Huge amount of data entry and frequent data manipulation(inserts, updates, deletes)
- Various but not costly SQLs
- Supports thousands of concurrent users
- We developed a new generic system (Harzemli) that produces surveys generic using DDL. Regional Offices use this program on the netbooks.

Institutional (Enterprise) Databases

- Institutional database has analysed, harmonized and controlled consistency eventual time series data.
- Only for internal users.
- User roles are defined for the statistical units on the newly created institutional databases.
- The statistical units can access only their own data using SAS codes , and just have select right.
- They can never delete, update, insert data by themselves.
- It is necessity for all departments to work in coordination with IT Department during the creation process of institutional databases.

Dissemination Databases

- Dissemination databases are the summary aggregated data repository which are created from institutional databases according to the information request
- The dissemination database is served on the TurkStat web page
- They are READ-ONLY system
- For external users, only confidential data is available
- It is free of charge, shown on web page
- Since it is open to whole world, supports many concurrent users

- There are aggregated time-series data about 62 subjects.
- End users can easily get desired data with the help of interface programs and dynamic reports
- The current web-based programmes are developed by our programmers with ZK framework and Oracle reports
- We are developing a new Central Dissemination Project that will show all dissemination data by one software programme

The main benefits of distinction to these 3 database is:

To maintain different business processes on different servers.

An advantage is performance. While we are making a bulk insertings or queries on institutional database, dissemination database doesn't effect these procedures. So the performance would not be decreased.

Another advantage is only aggregated data are served to internet users and access to micro data is **prevented**

Data Harmonization

The main subject of the Institutional databases is harmonization of different data

After careful investigations and analyses, it has been seen that our main problems comes from **using different classifications** from year to year and projects to projects

Such as

- Branch of economic activity
- Occupational group
- Educational status
- Marital Status
- Administrative Structure
- Dictionary tables

Data Warehouse in Turkstat

Being holded in our relational institutional database

Designed as star schema model

ROLAP (Relational OLAP) which means that all OLAP cubes are on relational database

Data Warehouse in Turkstat (Cont'd)

Being holded in our institution at the scope of European Union projects, Creating data warehouse process was started in 2006 , **with SAS**

And a prototype data warehouse consists of 1985, 1990 and 2000 population censuses were formed

Oracle BI suite is used.

Data Warehouse in Turkstat (Cont'd)

8 data warehouses / OLAP projects with 48 data marts:

- 2000 Population Census
- 1990 Population Census
- 1985 Population Census
- Labour Force Statistics
- Municipal Waste Statistics
- Municipal Water Statistics
- Foreign Trade Statistics
- Demographical Statistical

Datawarehouses in Institutional database. There are different schemas for each Datawarehouse.

SQL Procesures are used as ETL, to producing DW data.

Oracle BI Suite

OBIEE 11g is includes :

OBIEE Administration Tool

OBIEE Interface

BI Publisher

OBIEE Administration Tool

(Online) Oracle BI Administration Tool - 10.10.17.27

File Edit View Manage Tools Actions Window Help

Presentation

- Dış Adnks
- Dış Adnks Eğitim Medeni
- Dış Adnks Uygulama
- Sample Sales Lite
- Sample Targets Lite
- WEBHUKUMLU
- İç Nüfus 1985 Doğurganlık
- İç Nüfus 1985 Eğitim
- İç Nüfus 1985 Hanehalkı Konut
- İç Nüfus 1985 Tüm Nüfus
- İç Nüfus 1985 İstihdam
- İç Nüfus 1985 İşgücü
- İç Nüfus 2000 Doğurganlık
- İç Nüfus 2000 Eğitim
- İç Nüfus 2000 Hanehalkı Konut
- İç Nüfus 2000 Tüm Nüfus
- İç Nüfus 2000 İstihdam
- İç Nüfus 2000 İşgücü
- İç Nüfus1990 Doğurganlık
- İç Nüfus1990 Eğitim
- İç Nüfus1990 Hanehalkı Konut
- İç Nüfus1990 Hanehalkı Reisi
- İç Nüfus1990 Tüm Nüfus
- İç Nüfus1990 İstihdam
- İç Nüfus1990 İşgücü
- İç Yapısal İş İstatistikleri
- İç Çevre Belediye Atık Atık Tipi (Kaynak)
- İç Çevre Belediye Atık Belediye Nüfus Bilgileri
- İç Çevre Belediye Atık Bertaraf Miktarları
- İç Çevre Belediye Atık Tesis Bilgileri
- İç Çevre Belediye Atık Toplanan Atık Bilgileri
- İç Çevre Belediye Atık Tıbbi Atık Miktarları

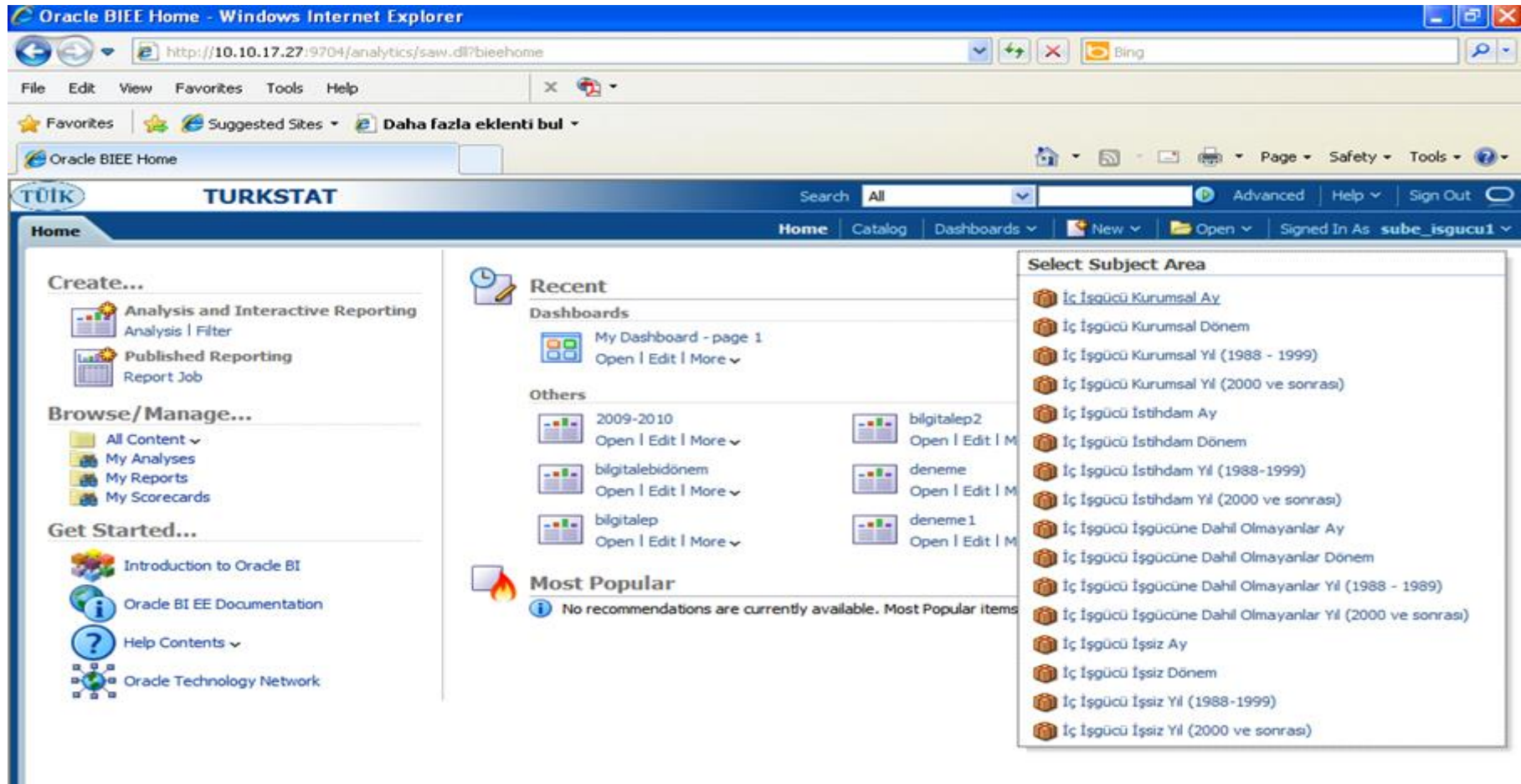
Business Model and Mapping

- İç Çevre Belediye Su İçme Suyu Hizmeti
- İç İşgücü Kurumsal Ay
- İç İşgücü Kurumsal Dönem - Yıl (1988 - 1999)
- İç İşgücü Kurumsal Yıl (2000 ve sonrası)
- İç İşgücü İstihdam Ay
- AY
- CİNSİYET
- EK İŞ OLUP OLMADIĞI
- EKSİK İSTİHDAM
- EĞİTİM DURUMU
- FİİLİ ÇALIŞMA SÜRESİ (SAAT)
- KENT - KIR
- MEDENİ DURUM
- MESLEK GRUBU (ISCO 88)
- REFERANS KİŞİYE YAKINLIK
- SOSYAL GÜVENLİK KURUMUNA KAYITLILIK
- SÜREKLİLİK DURUMU
- TARIM - TARIM DIŞI İSTİHDAM
- YAŞ GRUBU
- YETERSİZ İSTİHDAM
- YIL
- ZAMANA BAĞLI EKSİK İSTİHDAM
- ÇALIŞMA ŞEKLİ
- ÖLÇÜM İSTİHDAM AY
- İKTİSADİ FAALİYET KOLU (14 GRUP) - REV1
- İKTİSADİ FAALİYET KOLU (18 GRUP) - REV2
- İKTİSADİ FAALİYET KOLU (4 GRUP) - REV2
- İKTİSADİ FAALİYET KOLU (9 GRUP) - REV1
- İŞ ARAMA NEDENİ
- İŞTEKİ DURUM
- İŞYERİ BÜYÜKLÜĞÜ
- İŞYERİ DURUMU

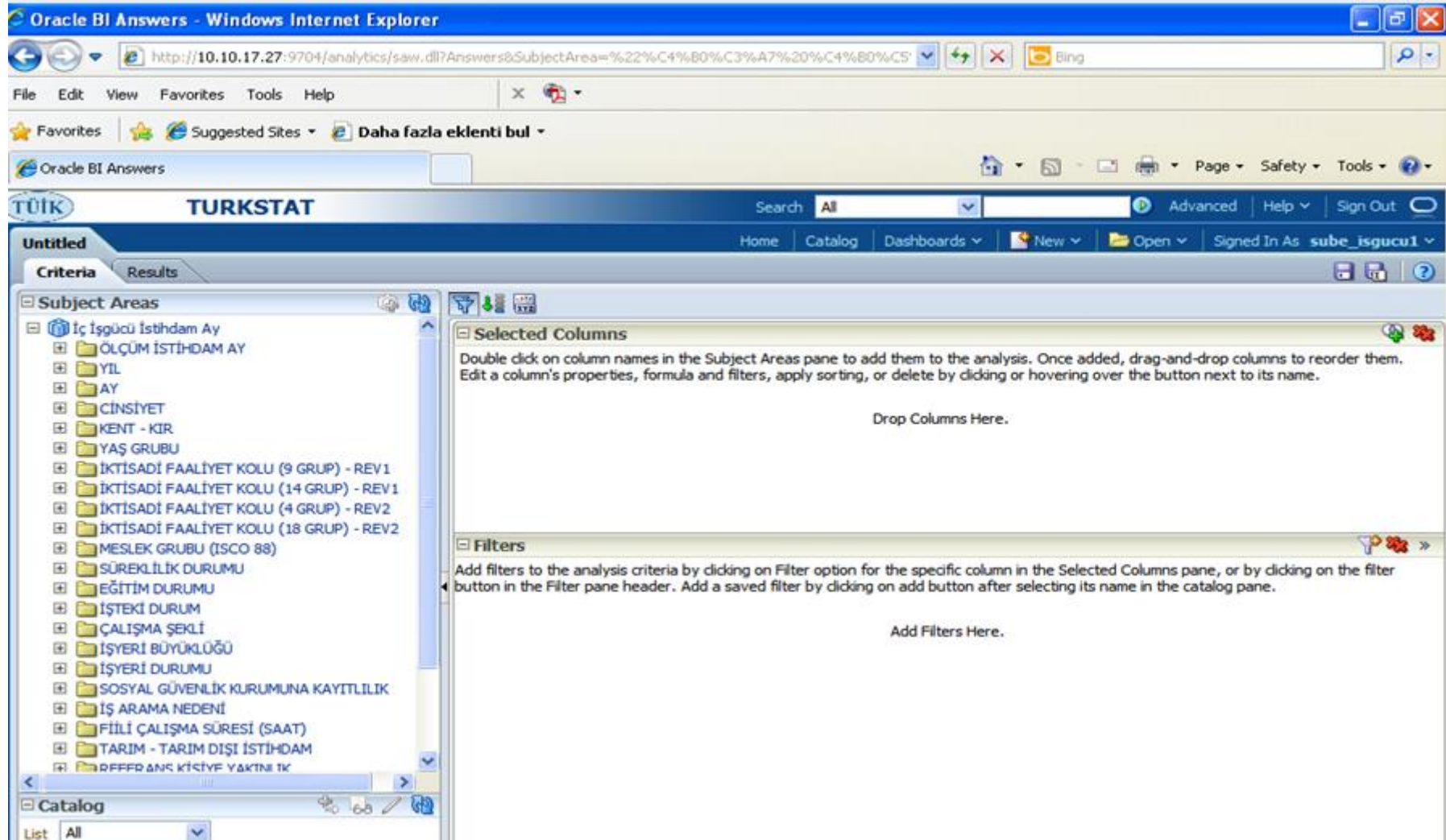
Physical

- dagitdb2
- kurumdb1
 - bi_kullanici
 - CEVRE_OLAP
 - BSU_DIM_GENELREFERANS
 - BSU_DIM_TAHSILATYERKODLARI
 - BSU_DIM_YIL
 - BSU_FACT_TABLO1
 - BSU_FACT_TABLO2
 - BSU_FACT_TABLO3
 - BSU_FACT_TABLO4
 - BSU_FACT_TABLO5
 - ABONE_SAYISI
 - BLD_KODU
 - BULTENNO
 - SATILAN_SU_MIKTARI
 - SU_SATIS_GELIRI
 - YER_KODU
 - YIL
- DENEME_DIM
- DENEME_FACT
- DIM_KA_ATIK_NITELIK
- DIM_KA_ATIK_TIPI
- DIM_KA_ATIK_YONTEM
- DIM_KA_BELEDIYE
- DIM_KA_BER_YONTEM
- DIM_KA_FORMDO_YIL
- DIM_KA_IL
- DIM_KA_MEVSIM
- DIM_KA_TESIS_FAAL_DURUMU
- DIM_KA_TESIS_TIPI
- DIM_KA_TESIS_YIL

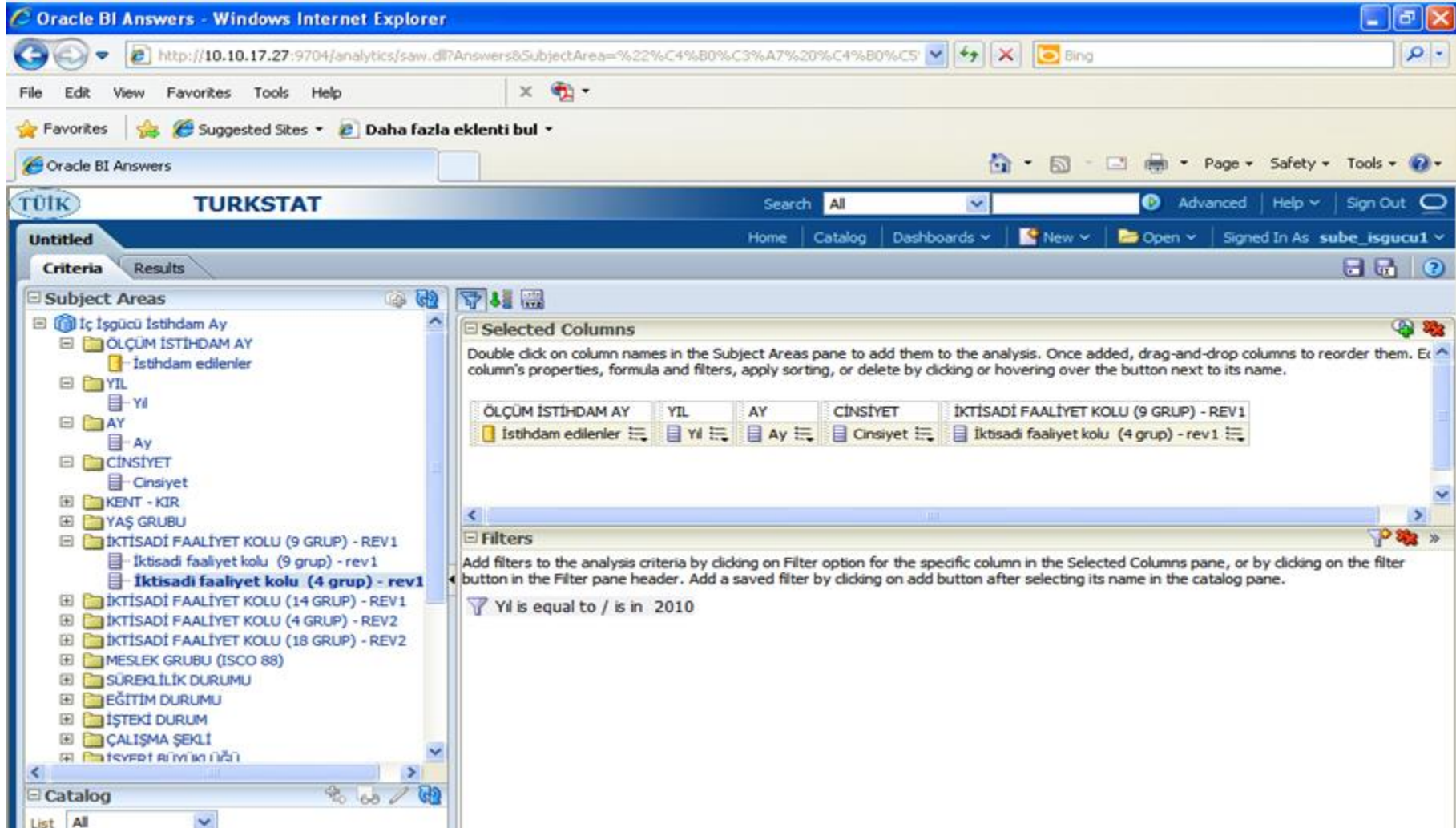
OBIEE Interface



OBIEE Interface



OBIEE Interface



The screenshot displays the Oracle BI Answers web interface within a Windows Internet Explorer browser window. The browser's address bar shows the URL: `http://10.10.17.27:9704/analytics/saw.dll?Answers&SubjectArea=%22%C4%B0%C3%A7%20%C4%B0%CS`. The page title is "Oracle BI Answers".

The main interface is titled "TURKSTAT" and includes a search bar with the text "All". The navigation menu includes "Home", "Catalog", "Dashboards", "New", "Open", and "Signed In As sube_igucu1".

The interface is divided into two main sections: "Criteria" and "Results". The "Criteria" section is currently active and contains the following panels:

- Subject Areas:** A tree view showing the hierarchy of data sources. The selected area is "İktisadi faaliyet kolu (4 grup) - rev1".
- Selected Columns:** A table listing the columns selected for the analysis. The columns are: "ÖLÇÜM İSTİHDAM AY", "YIL", "AY", "CİNSİYET", and "İKTİSADİ FAALİYET KOLU (9 GRUP) - REV1".
- Filters:** A section for adding filters to the analysis. The filter "Yıl is equal to / is in 2010" is currently applied.

The "Catalog" panel at the bottom shows a list of all available data sources.

OBIEE Interface (pivot table)

Oracle BI Answers - Windows Internet Explorer

http://10.10.17.27:9704/analytics/saw.dll?Answers&SubjectArea=%22%C4%B0%C3%A7%20%C4%B0%C5%

File Edit View Favorites Tools Help

Oracle BI Answers

TÜİK TURKSTAT

Search All Advanced Help Sign Out

Signed In As sube_igucu1

Untitled

Criteria Results

Subject Areas

- İç İlgücü İstihdam Ay
- ÖLÇÜM İSTİHDAM AY
- YIL
- AY
- CİNSİYET
- KENT - KIR
- YAŞ GRUBU
- İKTİSADİ FAALİYET KOLU (9 GRUP) - REV1
- İKTİSADİ FAALİYET KOLU (14 GRUP) - REV1
- İKTİSADİ FAALİYET KOLU (4 GRUP) - REV2
- İKTİSADİ FAALİYET KOLU (18 GRUP) - REV2
- MESLEK GRUBU (ISCO 88)
- SÜREKLİLİK DURUMU
- EĞİTİM DURUMU

Catalog

List All

Views

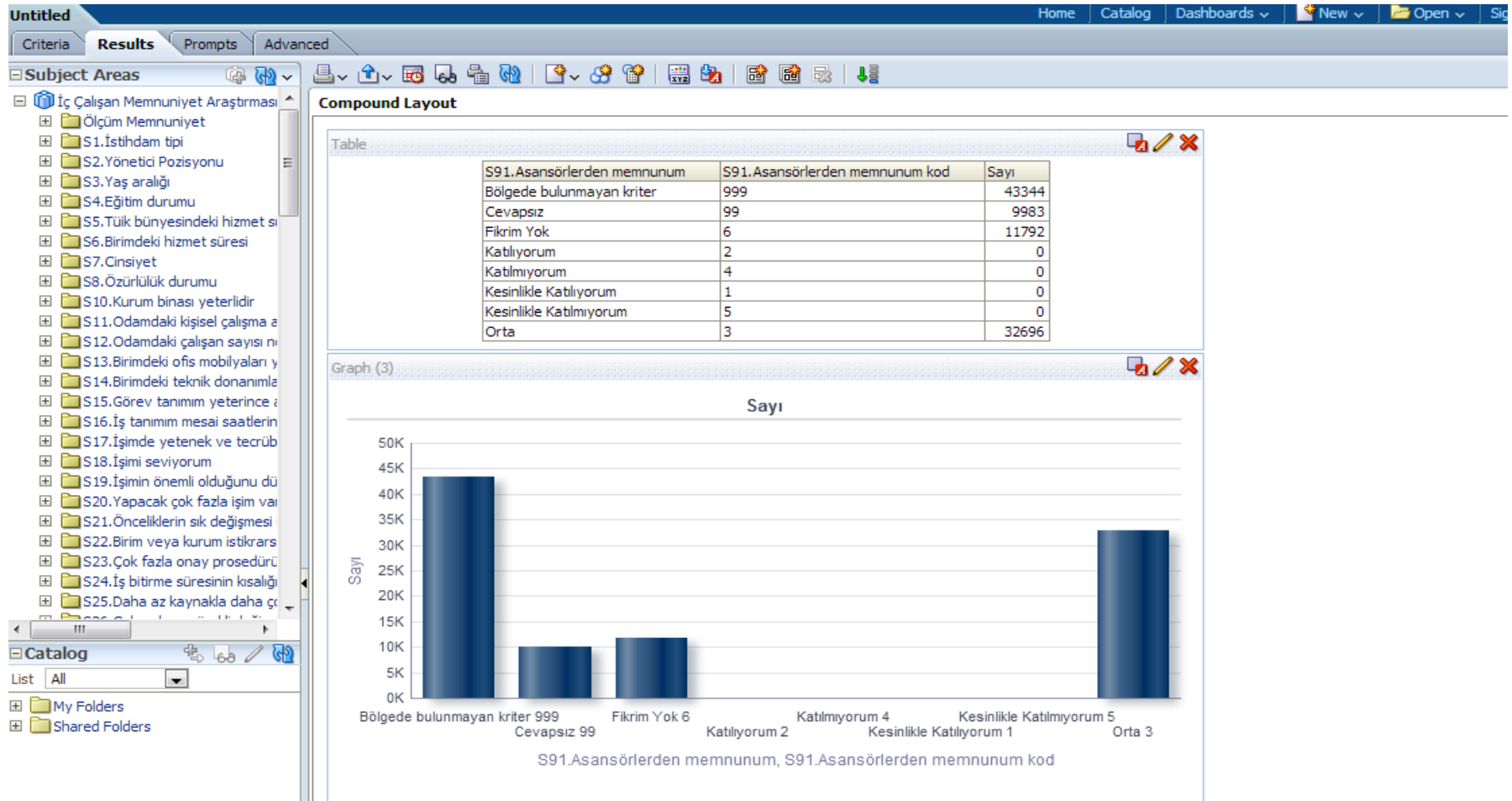
- Title
- Table
- Pivot Table

Compound Layout

Pivot Table

			Tarım	Sanayi	Ticaret	Hizmet
			İstihdam edilenler	İstihdam edilenler	İstihdam edilenler	İstihdam edilenler
Yıl	Ay	Cinsiyet				
2010	Ocak Total		5,039	5,441	4,283	6,398
	Ocak	Erkek	2,757	4,464	3,530	4,531
		Kadın	2,283	977	753	1,867
	Şubat Total		5,040	5,441	4,268	6,519
	Şubat	Erkek	2,771	4,435	3,517	4,624
		Kadın	2,269	1,006	750	1,895
	Mart Total		5,197	5,563	4,296	6,684
	Mart	Erkek	2,806	4,553	3,531	4,732
		Kadın	2,391	1,010	765	1,953
	Nisan Total		5,619	5,799	4,374	6,709
	Nisan	Erkek	2,932	4,787	3,597	4,746
		Kadın	2,688	1,012	777	1,962
Mayıs Total		5,831	5,986	4,459	6,779	
Mayıs	Erkek	2,947	4,976	3,644	4,806	
	Kadın	2,884	1,010	815	1,973	
Haziran Total		6,233	6,116	4,571	6,568	
Haziran	Erkek	3,099	5,066	3,734	4,707	
	Kadın	3,134	1,051	837	1,861	
Temmuz Total		6,284	6,208	4,540	6,446	

OBIEE Interface (table and chart)



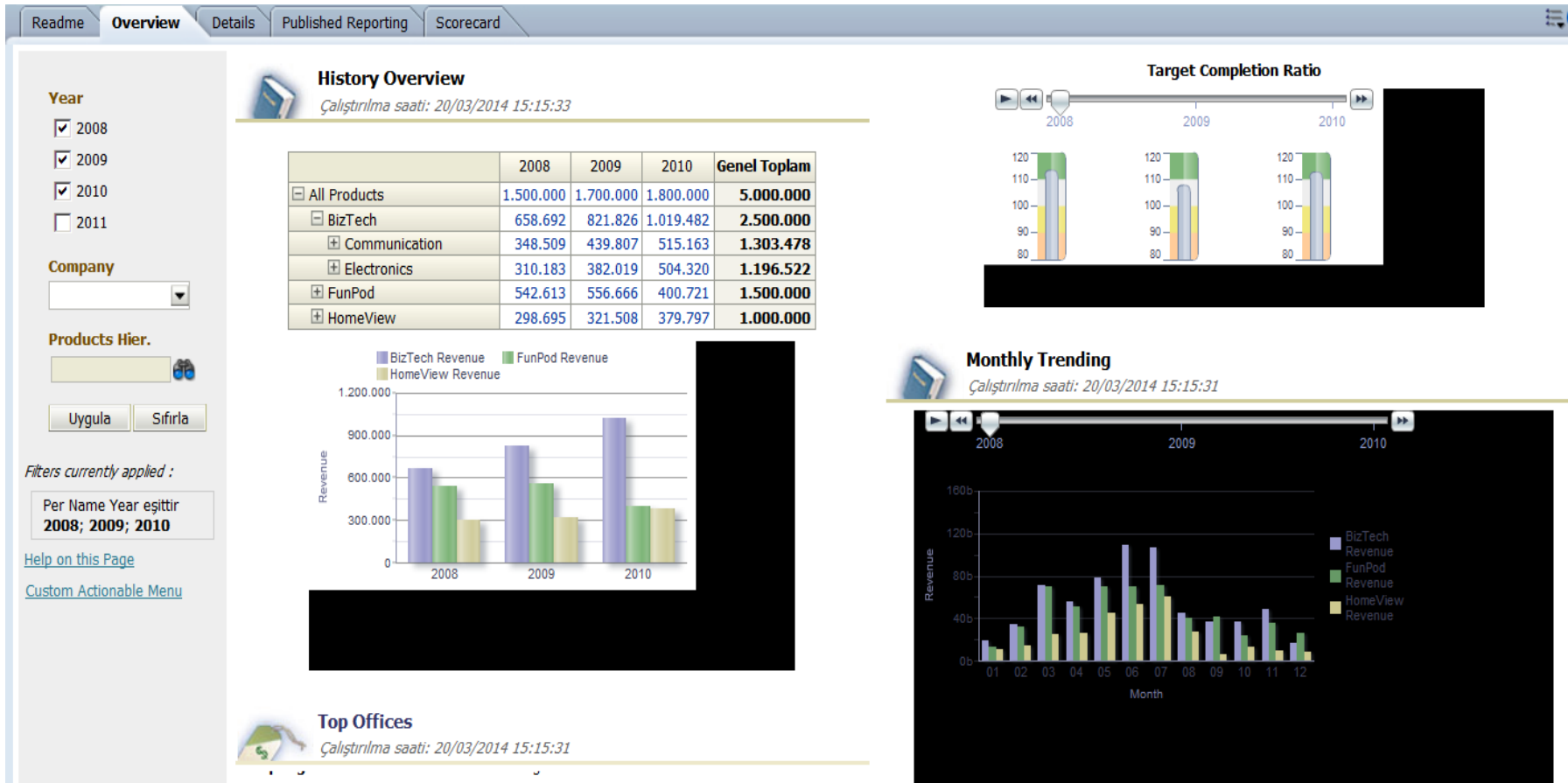
OBIEE Interactive Dashboards

Simply containers for reports and other content (like chart, maps..)

Allows multiple reports to be displayed in a tabbed interface

Used to provide highly summarized and graphical content

An Example Interactive Dashboard



Summary

- ✓ Databases of the TurkStat are Production, Institutional and Dissemination
- ✓ Separating databases to 3 database is useful both performance, management, confidentiality
- ✓ We have some Data Warehouse Projects developed by Oracle BI Suite
- ✓ Oracle BI suite has Administration Console, Presentation Interface, Reports Building tool
- ✓ Dashboards are containers that can include various and dynamic reports, maps, charts etc., and give very easy and valuable data to the users