

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

مدیریت منابع داده ای

مدرس: دکتر مسعود یقینی

پائیز ۱۳۸۸

اصطلاحات و مفاهیم سازماندهی فایل

اصطلاحات و مفاهیم سازماندهی فایل

- **بیت (Bit):** کوچکترین واحد داده است؛ یک رقم باینری که می تواند 0 یا 1 باشد.
- **بایت (Byte):** گروهی از بیتها که نمایانگر یک کاراکتر است.
- **فیلد (Field):** گروهی از لغات یا اعداد کامپیوتری
- **رکورد (Record):** گروهی از فیلدهای مرتبط
- **فایل (File):** گروهی از رکوردهای هم نوع
- **پایگاه داده (Database):** گروهی از فایل‌های مرتبط

اصطلاحات و مفاهیم سازماندهی فایل

سلسله مراتب داده

سلسله مراتب

پایگاه داده

فایل

رکورد

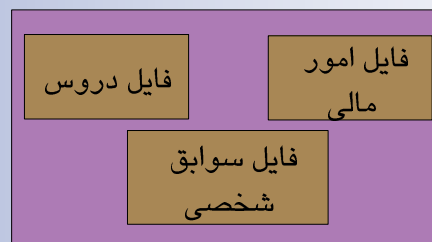
فیلد

بایت

بیت

مثال

پایگاه داده دانشجو



فایل دروس

| نام | درس | تاریخ | نمره |
|--------------|--------|------------|------|
| جان استوارت | IS 101 | پائیز 2003 | B+ |
| کارن تیلور | IS 101 | پائیز 2003 | A |
| امیلی وینسنت | IS 101 | پائیز 2003 | C |

| نام | درس | تاریخ | نمره |
|-------------|--------|------------|------|
| جان استوارت | IS 101 | پائیز 2003 | B+ |

جان استوارت (فیلد نام)

01001010 حرف J بر اساس کد ASCII

0

اصطلاحات و مفاهیم سازماندهی فایل

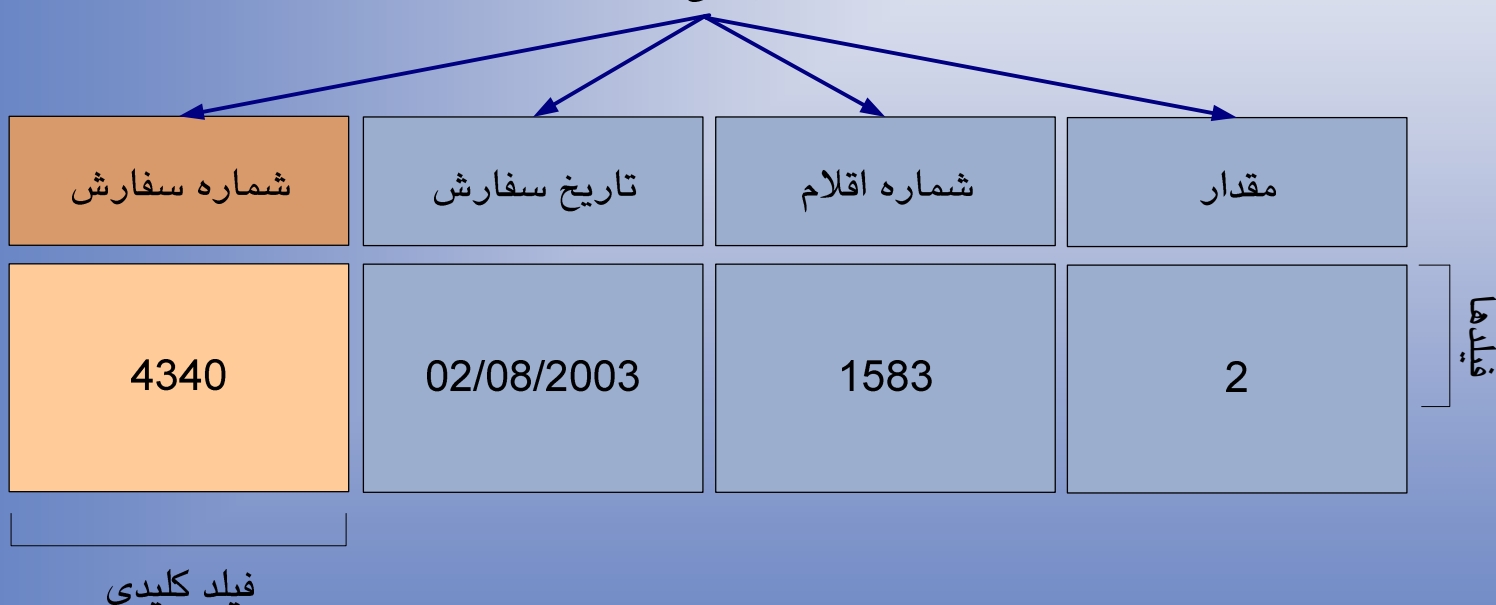
- **موجودیت (Entity):** فرد، مکان، شیء، یا رویدادی که درباره آن اطلاعات جمع آوری می شود.
- **ویژگی (Attribute):** قطعه ای از اطلاعات که یک موجودیت بخصوص را توصیف می کند.
- **فیلد کلیدی (Key Field):** فیلدی منحصر به فردی که هر رکورد از یک فایل را شناسایی می کند.

اصطلاحات و مفاهیم سازماندهی فایل

موجودیتها و ویژگیها

موجودیت : سفارش

ویژگی ها



رویکرد پایگاه داده به
مدیریت داده

رویکرد پایگاه داده به مدیریت داده

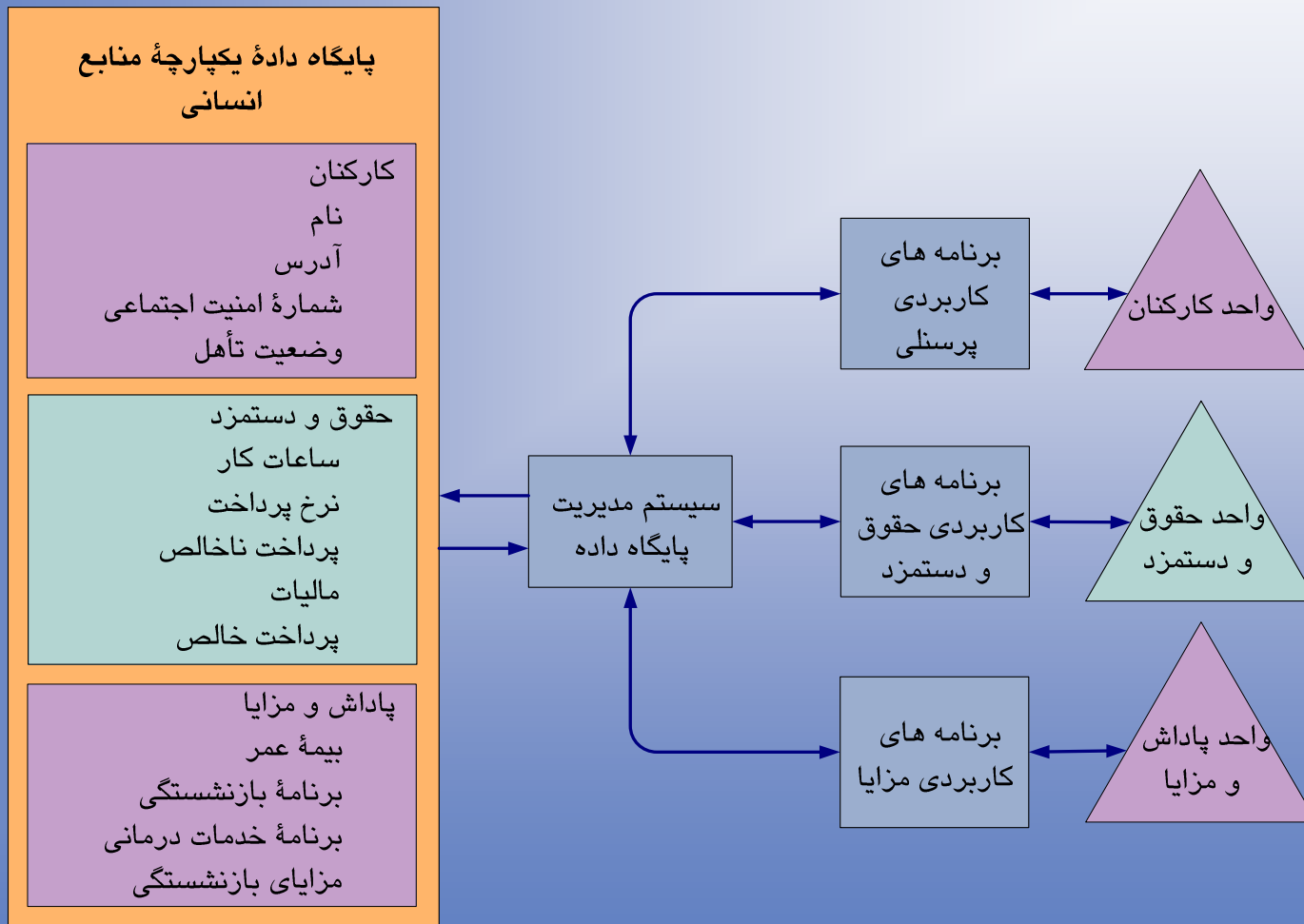
پایگاه داده

Database

- مجموعه ای است از داده های متمرکز.
- افزودن داده را کنترل می کند.
- داده ها ذخیره شده اند تا برای کاربر در یک مکان واحد جلوه گر شوند.
- به چندین سیستم کاربردی خدمت رسانی می کند.

رویکرد پایگاه داده به مدیریت داده

محیط یک پایگاه داده



رویکرد پایگاه داده به مدیریت داده

سیستم مدیریت پایگاه داده

Database Management System (DBMS)

- پایگاه های داده را ایجاد و از آنها نگهداری می کند.
- نیاز به تعریف داده را از بین می برد.
- حکم واسطی بین برنامه های کاربردی و فایل های داده های فیزیکی را دارد.
- دیدهای فیزیکی و منطقی داده را از هم جدا می کند.

رویکرد پایگاه داده به مدیریت داده

مزایای استفاده از DBMS:

- کاهش افزونگی داده
- حذف عدم سازگاری داده
- حذف وابستگی برنامه به داده
- افزایش دسترسی و دسترس پذیری داده
- فراهم آوردن امکان مدیریت متمرکز داده، استفاده از داده، و امنیت

رویکرد پایگاه داده به مدیریت داده

DBMS رابطه ای Relational DBMS

- داده را در قالب جداول دو بعدی نمایش می دهد.
- داده را از طرق جداولی بر مبنای عناصر داده ای مشترک به هم مرتبط می سازد.
- مثال:
- Oracle, DB2 برای سیستمهای بزرگ
- MS SQL Server برای سیستمهای حد متوسط
- Microsoft Access و Oracle Lite برای سیستمهای کوچک

رویکرد پایگاه داده به مدیریت داده

مدل داده رابطه ای



رویکرد پایگاه داده به مدیریت داده

سه عملیات پایه ای در یک پایگاه داده رابطه ای

- **انتخاب (Select):** زیر مجموعه ای از ردیفها ایجاد می کند که معیارهای بخصوصی را برآورده می سازد.
- **ادغام (Join):** جداول رابطه ای را برای تأمین کاربران با اطلاعات مناسب، با هم ترکیب می کند.
- **طرح ریزی (Project):** کاربران را قادر به ایجاد جداول جدیدی می سازد که تنها حاوی اطلاعات مناسب هستند.

رویکرد پایگاه داده به مدیریت داده

سه عملیات پایه ای از یک DBMS رابطه ای

سفارش

| شماره سفارش | تاریخ سفارش | تاریخ تحویل | شماره قطعه | تعداد سفارش |
|-------------|-------------|-------------|------------|-------------|
| 1634 | 02/02/04 | 02/22/04 | 152 | 2 |
| 1635 | 02/12/04 | 02/28/04 | 137 | 3 |
| 1636 | 02/13/04 | 03/01/04 | 145 | 1 |

قطعه

| شماره قطعه | شرح قطعه | قیمت واحد | شماره تأمین کننده |
|------------|-------------|-----------|-------------------|
| 137 | قفل درب | 22.50 | 4058 |
| 145 | دستگیره درب | 26.25 | 2038 |
| 150 | درزگیر درب | 6.00 | 4058 |
| 152 | کمپرسور | 70.00 | 1125 |

تأمین کننده

| نام تأمین کننده | نام تأمین کننده | آدرس تأمین کننده |
|-----------------|-----------------|------------------|
| 4058 | شرکت CBM | |
| 2038 | شرکت Ace | |
| 1125 | شرکت Bryant | |

انتخاب شماره قطعه = 137 یا 152

ادغام از طریق شماره تأمین کننده

| شماره قطعه | شماره تأمین کننده | نام تأمین کننده | آدرس تأمین کننده |
|------------|-------------------|-----------------|------------------|
| 137 | 4058 | شرکت CBM | |
| 152 | 1125 | شرکت Bryan | |

ستونهای انتخاب پروژه

ایجاد یک محیط
پایگاه داده

ایجاد یک محیط پایگاه داده

دو فعالیت در طراحی یک پایگاه داده

- طراحی مفهومی یا منطقی (Conceptual or Logical Design):

مدل خلاصه شده ای است از پایگاه داده از زاویه دید کسب و کار.

- طراحی فیزیکی (Physical Design): مدلی که نشان می دهد

پایگاه داده چگونه بر روی تجهیزات ذخیره سازی، سازماندهی شده است.

ایجاد یک محیط پایگاه داده

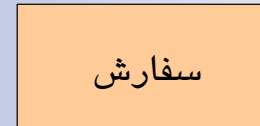
طراحی مفهومی پایگاه داده Conceptual Database Design

- رابطه بین عناصر داده ای را شناسایی می کند.
- کاراترین راه گروه بندی عناصر داده ای را شناسایی می کند.
- عناصر داده ای تکراری را شناسایی می کند.
- گروه های داده ای مورد نیاز سیستمهای کاربردی بخصوصی را شناسایی می کند.
- **نمودار رابطه موجودیت (Entity-Relationship Diagram)** متدولوژی برای مستند سازی پایگاه های داده که رابطه بین موجودیتهای مختلف پایگاه داده را نشان می دهد.

ایجاد یک محیط پایگاه داده

یک نمودار رابطه موجودیتها

ویژگی های موجودیت
شماره سفارش
تاریخ سفارش
تاریخ تحویل
شماره قطعه
تعداد قطعه

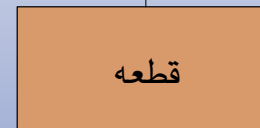


M



M

شماره قطعه
شرح قطعه
قیمت واحد
شماره تأمین کننده

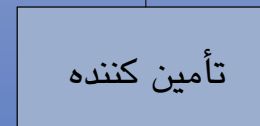


M



1

شماره تأمین کننده
نام تأمین کننده
آدرس تأمین کننده



ایجاد یک محیط پایگاه داده

نرمال سازی

(Normalization)

• فرآیند ایجاد ساختارهای داده کوچک، پایدار، و در عین حال منعطف از گروههای پیچیده داده، زمانی که یک پایگاه داده ای رابطه ای در حال طراحی است.

ایجاد یک محیط پایگاه داده

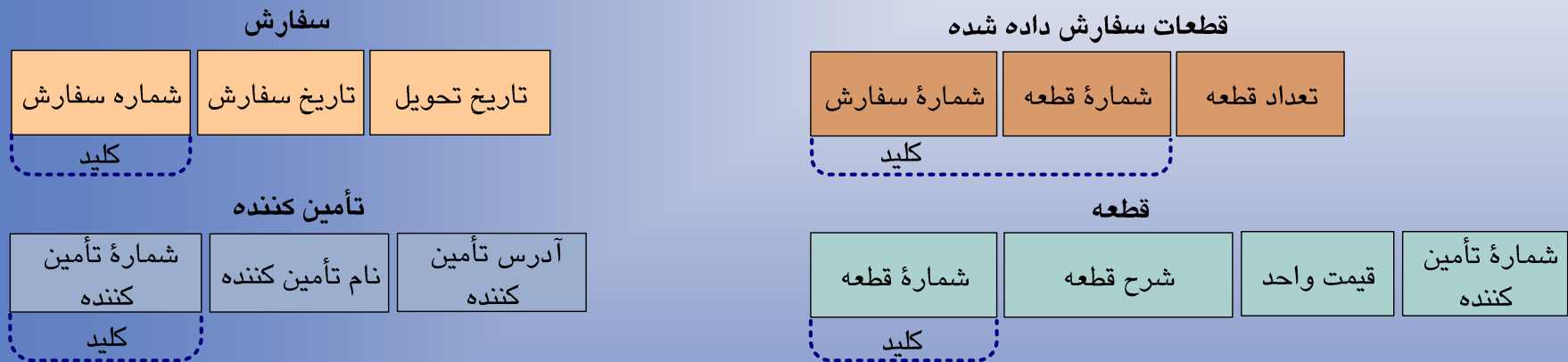
یک رابطه غیر نرمال برای موجودیت سفارش

سفارش

| | | | | | | | | | |
|-------------|------------|------------|----------|-----------|-------------------|-----------------|------------------|-------------|-------------|
| شماره سفارش | تعداد قطعه | شماره قطعه | شرح قطعه | قیمت واحد | شماره تأمین کننده | نام تأمین کننده | آدرس تأمین کننده | تاریخ سفارش | تاریخ تحویل |
|-------------|------------|------------|----------|-----------|-------------------|-----------------|------------------|-------------|-------------|

ایجاد یک محیط پایگاه داده

یک رابطه نرمال شده برای موجودیت سفارش



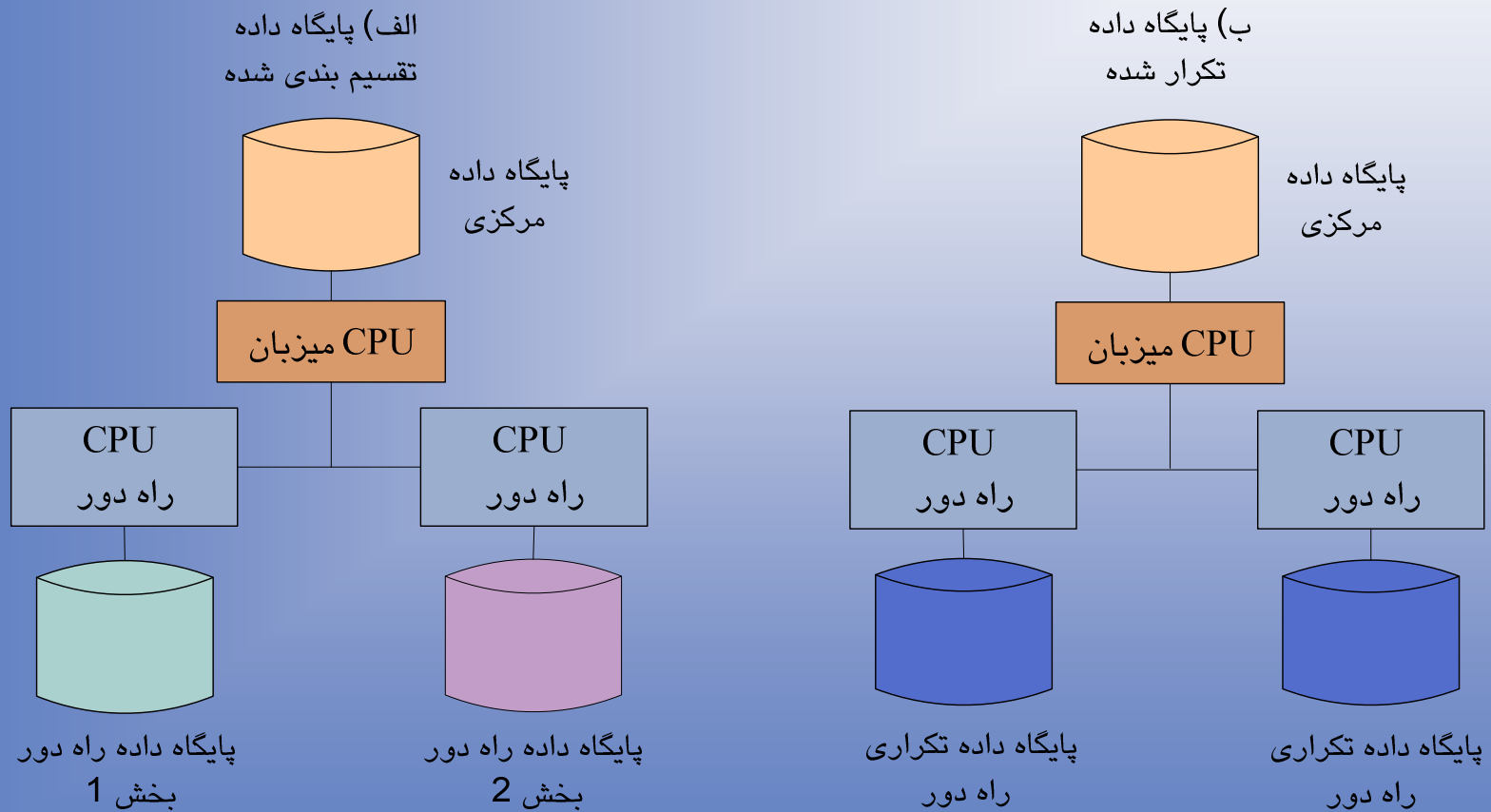
ایجاد یک محیط پایگاه داده

پایگاه داده توزیع شده

- پایگاه داده توزیع شده (**Distributed Database**): پایگاه داده ای است که در بیش از یک مکان فیزیکی واقع شده باشد؛ به بیش از یک مکان تقسیم بندی یا تکرار شده است.

ایجاد یک محیط پایگاه داده

پایگاه های داده توزیع شده



ایجاد یک محیط پایگاه داده

پایگاه داده توزیع شده

مزایا:

- افزایش خدمت رسانی و پاسخگویی به کاربران محلی
- کاهش آسیب پذیری به یک سایت واحد مرکزی بزرگ

معایب:

- وابستگی به خطوط ارتباط از راه دور
- ریسک امنیتی از طریق توزیع داده های حساس
- به روز رسانی داده های مرکزی با داده های محلی

روند پیشرفت
پایگاه های داده

روند پیشرفت پایگاه های داده

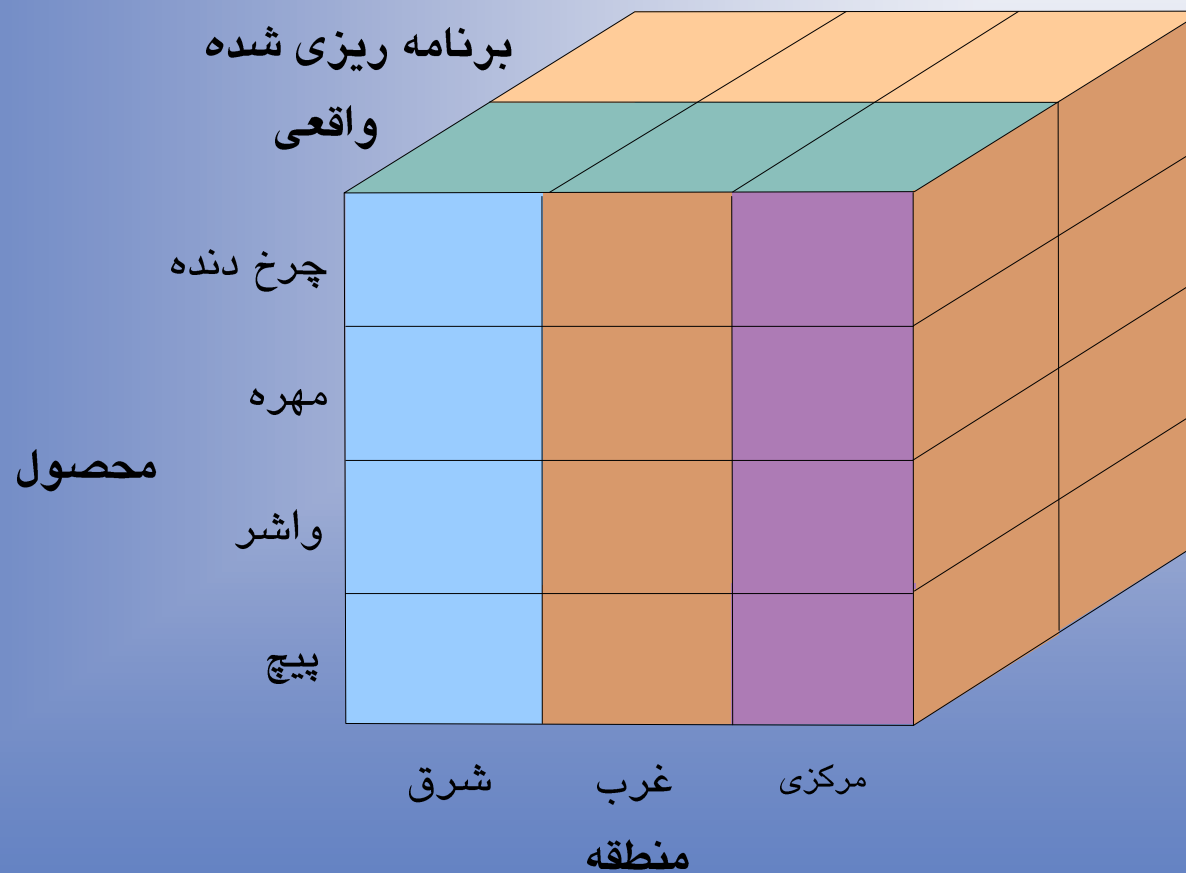
پردازش تحلیلی آنلاین

Online Analytical Processing (OLAP)

- OLAP کاربران را قادر می سازد داده های واحد را به شکل های مختلف با استفاده از ابعاد چندگانه ملاحظه کنند.
- هر جنبه از اطلاعات - محصول، قیمت، منطقه - نمایانگر بعد متفاوتی است.

روند پیشرفت پایگاه های داده

مدل چند بعدی داده



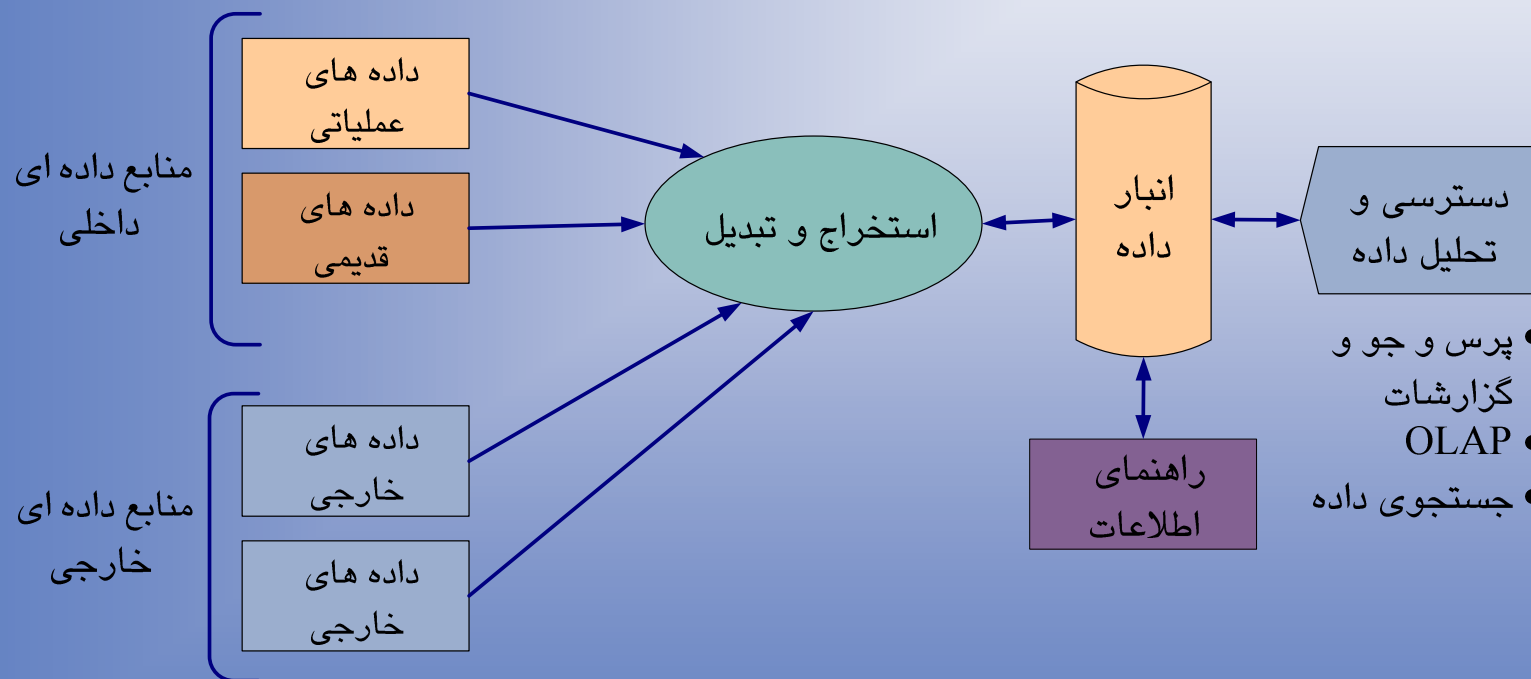
روند پیشرفت پایگاه های داده

● **انبار داده (Data Warehouse):** داده های جاری و قدیمی را برای گزارش دهی و تحلیل ذخیره می سازد.

● **بازار داده (Data Mart):** زیر مجموعه ای است از انبارهای داده با خلاصه ای از داده برای کاربران خاص.

روند پیشرفت پایگاه های داده

اجزای یک انبار داده



روند پیشرفت پایگاه های داده

مزایای انبارهای داده

- بهبود اطلاعات و قابلیت دسترسی به آنها
- قابلیت مدل کردن و مدلسازی مجدد داده
- قادر ساختن دسترسی به داده بدون تأثیر گذاری بر عملکرد سیستم های عملیاتی موروثی زیربنایی سازمان

روند پیشرفت پایگاه های داده

پایگاه داده فرارسانه ای Hypermedia Database

- داده را در قالب شبکه ای از نودها سازماندهی می کند.
- نودها را در قالب الگوهای مشخص شده توسط کاربر به هم مرتبط می سازد.
- از متن، گرافیک، صدا، ویدئو، و برنامه های اجرایی پشتیبانی می کند.

روند پیشرفت پایگاه های داده

داده کاوی

● **داده کاوی (Data Mining):** تکنیکهایی برای یافتن الگوها و روابط پنهان در حجم عظیمی از داده برای استنتاج قوانینی برای پیش بینی روندهای آینده.

روند پیشرفت پایگاه های داده

نمونه هایی از کاربرد داده کاوی

● **Disco S.A.**: فروشگاه های زنجیره ای آرژانتین از استخراج داده برای تحلیل الگوهای خرید بیش از ۵/۱ میلیون مشتری که دارای برنامه خرید متناوبی در بیش از ۲۰۰ فروشگاه هستند، استفاده می کند.

● **رستوران Red Robin**: این رستورانها از استخراج داده برای تحلیل منوهای ۸۷ رستوران خود استفاده می کند. با تحلیل قیمت، هزینه، و کیفیت هر یک از موادی که در غذاهای سرو شده به کار می رود، Red Robin می تواند اقلام منو را که هم دارای تقاضای زیاد و هم سود حاشیه ای بالاست، شناسایی کند.

مدیریت منابع داده ای