



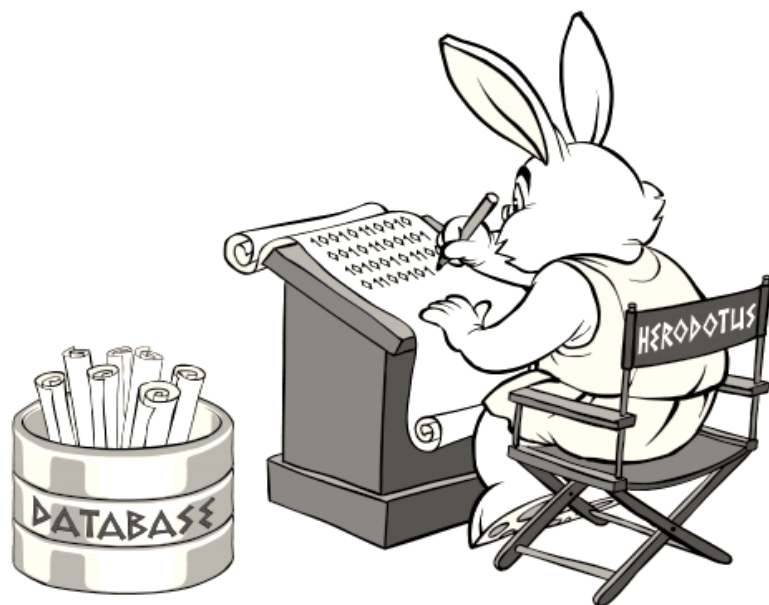
# پایگاه داده‌ها

مدرس: میلاد وزان

دانشگاه شهید بهشتی - دانشکده ریاضی - گره آموزشی آمار

<https://dbsbu.github.io>

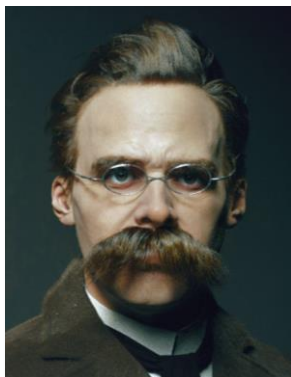
## فصل ۱: مقدمه



نیم سال دوم ۱۴۰۳-۱۴۰۴

"آن که همیشه شاگرد می ماند، آموزگارِ خویش را پاداشی به سزا نمی دهد."

فردریش نیچه | چنین گفت زرتشت



## تعاریف

- داده
- تعریف: داده به حقایق خام گفته می‌شود که به تنهایی معنای خاصی ندارند، اما با پردازش و تحلیل می‌توانند دانش و اطلاعات مفید تبدیل شود.
  - درک این حقایق، در شکل خام خود، بدون پردازش اضافی، می‌تواند دشوار باشد.
  - خصوصیات: به خورده‌ی خودفاقد زمینه و مفهوم است.
  - مثال: 169.3, 164.7, 166.4, 175.3-161.2

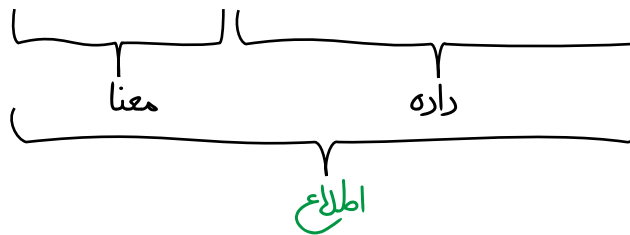
💡 داده‌ها حقایق خام هستند.

کلمه‌ی خام نشان می‌دهد که حقایق هنوز پردازش نشده‌اند تا اطلاعاتی را آشکار کنند.

- اطلاع
- تعریف: اطلاع زمانی ایجاد می‌شود که داده‌ها پردازش، سازماندهی و ساختار بندی شوند تا زمینه و معنا را ارائه دهند.
  - اطلاع اساساً داده‌های پردازش شده است.
  - خصوصیات: معنی و مفهوم دارد.
  - مثال: 169.3, 164.7, 166.4, 175.3-161.2 قد دانش آموزان ۱۵ ساله است.

اطلاع = اختصاص معنا + داده

از دید پایگاه داده، این اختصاص معنا در واقع همان پردازش است و این پردازش از طریق کوئری پرس‌وجو یا پرس‌وجو انجام می‌شود.



مثال‌های بیشتر

- مثال ۱:
  - داده: ۱۰۰-۵۹-۳۰۰
  - معنا: تعداد فروشنده ماهانه محصولات در سه ماه مختلف
- مثال ۲:
  - داده: ۱۵-۱۷-۱۹-۱۸-۱۶-۱۹
  - معنا: نمرات آزمون پایگاه داده دانشجویان گروه آمار
- مثال ۳:
  - داده: ۵۰۰-۵۵-۷۰۰
  - معنا: مصرف برق ماهانه یک خانواده در سه ماه مختلف
- مثال ۴:
  - داده: ۲۵۶۵
  - معنا: تعداد کتاب‌های موجود در کتابخانه ریاضی دانشگاه شهید بهشتی
- مثال ۵:
  - داده: ۱۲۷
  - معنا: تعداد مقالات منتشر شده توسط اعضای هیئت علمی دانشکده ریاضی در سال گذشته

## یک جدول در یک پایگاه داده در درجه اول یک ظرف یک مکان ذخیره، برای داده است.

- حقایق و ارقام خام را بدون هیچ معنا یا زمینه ذاتی ذخیره می‌کند.
- داده‌های ذخیره‌شده در جداول زمانی به اطلاعات تبدیل می‌شوند که از طریق پرس و جوها پردازش، تجزیه و تحلیل یا تبدیل شوند تا بینش یا پاسخی به سؤالات خاص ارائه دهد.
- به عنوان مثال:
  - جدولی حاوی نام و سن دانشجویان فقط داده‌های خام را ذخیره می‌کند.
  - اجرای یک پرس و جو برای یافتن میانگین سنی دانشجویان این داده‌ها را به اطلاعات تبدیل می‌کند. بنابراین، در حالی که جداول داده‌ها را نگه می‌دارند، فرآیند پرس و جو از این جداول آن داده‌ها را به اطلاعات تبدیل می‌کند.

😊 جداول داده‌ها را ذخیره می‌کنند، اما جداول در واقع کاری انجام نمی‌دهند.  
 جدول‌ها اسم هستند ← 'تبدیلی انجام نمی‌دهد، صرفاً ذخیره داده'  
 کوئری‌ها فعل هستند ← 'تبدیل داده با اطلاع'

### داده‌ها

نمبره	نام درس	نام و نام خانوادگی دانشجو	شماره دانشجویی
۱۲	پایگاه داده	میلاد وزات	۱۵۵۱
۱۸	پایگاه داده	امیرحسین پارسا	۱۵۵۲
۱۷	پایگاه داده	علی رضا جهان بخش	۱۵۵۳
۱۹	پایگاه داده	آرتین رضانی	۱۵۵۴

### اطلاع

نمبره	شماره دانشجویی
۱۸	۱۵۵۲
۱۹	۱۵۵۴

تبدیل داده به اطلاع

نمبره و شماره دانشجویی دانشجویانی که نمره درس پایگاه داده را بالاتر از ۸ اکسب کرده‌اند.

گزارش 'کوئری'



**نکته:** یک جدول هم‌چنین می‌تواند به عنوان اطلاعات در نظر گرفته شود؛ 'چه زمانی؟؟؟؟' زمانی که نتیجه یک کوئری را نمایش دهد.

- تعریف:** دانش، اطلاعاتی است که تحت تجزیه و تحلیل بیشتر قرار گرفته است و به درک و بینش عمیق‌تری منجر می‌شود.
- دانش با افزودن تجربه، تفسیر و قضاوت بر روی اطلاعات ایجاد می‌شود و به آن اجازه می‌دهد تا برای حل مسائل یا ایجاد راه‌حل‌های نوآورانه به کار رود.
- خصوصیات:** دانش برای هر فردی منحصر به فرد است و انباشته‌ای از تجربه و بینش گذشته است.
- مثال:** قریب‌ترین دانش آموز ۱۷۵.۳ سانتی متر است.

کشف دانش از پایگاه داده‌ها در واقع فرآیند تشخیص الگوهای جالب، معتبر، بالقوه مفید و کاملاً قابل فهم از داده‌ها است.

این الگوها معمولاً در میان حجم عظیمی از داده‌ها مخفی هستند. 😊

▪ مثال - پایگاه داده اطلاعات بیماران مبتلا به سکت قلبی:

▪ جالب:

- یافتن ارتباط بین عوامل خطر ساز مانند فشار خون بالا، دیابت، و چربی خون بالا با بروز سکت قلبی
- شناسایی گروه‌های پرخطر برای ابتلا به بیماری قلبی خاص
- شناسایی رن‌های خاص که با افزایش ریسک ابتلا به بیماری‌های قلبی مرتبط هستند.
- تشخیص زودهنگام ناهنجاری‌ها در سیگنال‌های نوار قلب که ممکن است به سکت قلبی منجر شود
- معتبر: تأیید صحت این الگوها با استفاده از داده‌های گذشته‌ای بیماران
- بالقوه مفید: برنامه ریزی پیگیری‌ها برای کاهش ریسک ابتلا به بیماری قلبی
- قابل فهم: ارائه نمودارهایی که تغییرات عوامل خطر ساز را طی زمان نشان دهند.

بسیاری کشف دانش را معادل با داده‌کاوی در نظر می‌گیرند، چرا که داده‌کاوی در واقع به معنی استخراج دانش از پایگاه داده‌هاست.

سیلبرشاتر

اهداف داده‌کاوی

- پیش‌بینی: برای پیش‌بینی وضعیت احتمالی آینده بر اساس رویدادهای قبلی.
- با توجه به داده‌های فروش سال‌های گذشته:
- آیا می‌توانیم پیش‌بینی کنیم که برای فصل آینده چه مقدار کالا در انبار باید داشته باشیم؟
- توضیحات: دلیل وقوع برخی رویدادها چیست؟
- دلایل فروش بهتر خودروهایی که تولیدکننده نسبت به محصولات مشابه سایر تولیدکنندگان چیست؟
- تشخیص استثنا: ممکن است موقعیت‌هایی «سوابق» در پایگاه داده ما وجود داشته باشد که با چیزی غیرمادی مطابقت دارد.
- آیا می‌توانیم شکایات مشتریان را که نادرست یا همراه کننده هستند تشخیص دهیم؟
- آیا می‌توان گزارش‌های مالی را که برای اهداف متقلبانه دستکاری شده اند کشف کرد؟
- آیا می‌توان ادعاهای بیمه‌ای را که در واقع کلاهبرداری هستند شناسایی کرد؟

انواع مسائل داده‌کاوی / تحلیل داده

- توضیحاتی: چه اتفاقی افتاده است؟
- تشخیصی: چرا این اتفاق افتاد؟
- پیش‌بینی: چه اتفاقی خواهد افتاد؟
- تجویزی: در مورد آن چه باید کرد؟

پایگاه داده

فصل ۱: مقدمه

مجموعه‌ای است از داده‌های ذخیره شده، پایا، مجتمع و مبتنی بر یک ساختار، بهم مرتبط، تا حد ممکن فاقد افزونگی، تحت کنترل یک سیستم

متمرکز، مورد استفاده یک یا چند کاربر در یک سازمان «در یک محیط عملیاتی» به طور اشتراکی و هم‌روند

هر محیطی که برای آن سیستم ذخیره داده و بازیابی اطلاعات طراحی شود. --- محیط عملیاتی بیمارستان، دانشگاه و ...

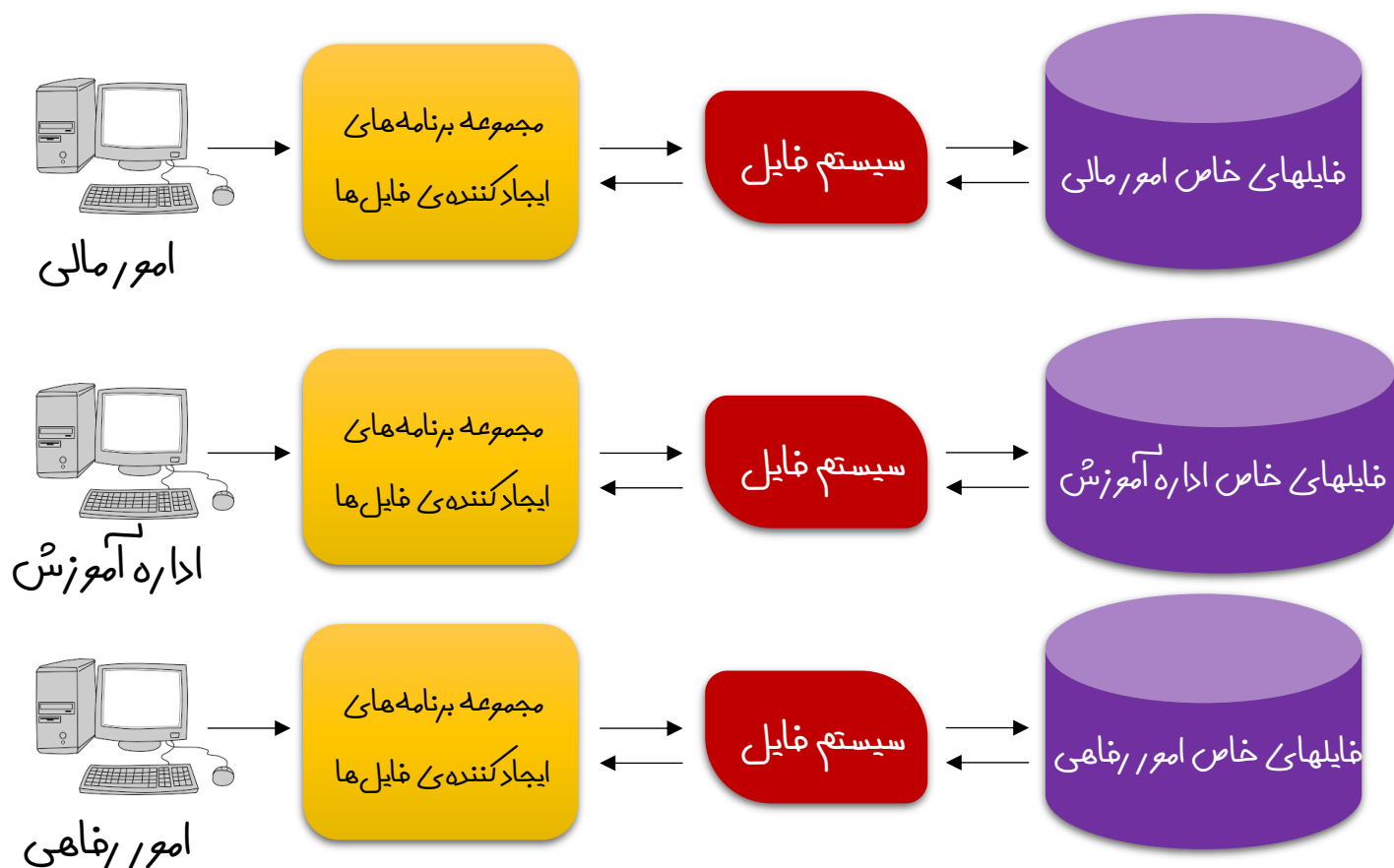
## انواع سیستم‌های ذخیره و بازیابی

- رویکرد فایلینگ «نپایگاهی»
- رویکرد پایگاهی «بانکی»

### رویکرد فایلینگ

در محیط عملیاتی دانشگاه که شامل بخش‌های مختلف است، سه بخش اصلی یعنی امور آموزش، امور رفاهی و امور مالی را در نظر می‌گیریم. هدف ما ایجاد یک سیستم ذخیره و بازیابی کامپیوتری برای این بخش‌ها است. فرض بر این است که تنها موجودیت مورد نظر در این سیستم، دانشجو باشد و هر سه بخش مذکور نیاز به اطلاعات مربوط به این موجودیت دارند.

در روش فایلینگ، هر زیرمحیط عملیاتی به طور مستقل مورد بررسی قرار می‌گیرد و برای هر زیر مجموعه، یک سیستم خاص طراحی و تولید می‌شود که تنها پاسخگوی نیازهای همان زیرمحیط باشد. با توجه به مثال دانشگاه دارای یک سیستم کاربردی منحصر به فرد و جداگانه خواهد بود.





## معایب رویکرد فایلینگ

- افزودگی داده‌ها و ناسازگاری: افزودگی داده‌ها به معنای ذخیره شدن یک داده در چندین محل مختلف است. این پدیده نه تنها باعث مصرف فضای اضافی بر روی دیسک می‌شود، بلکه در صورت نیاز به تغییر بخشی از داده، باید آن تغییر را در تمام جاهایی که آن داده ذخیره شده اعمال کرد. این امر منجر به ناسازگاری بین نسخه‌های مختلف همان داده می‌شود. بدین معنی که رونوشت‌های مختلف از یک داده مشابه، با یکدیگر سازگار نخواهند بود. **به عنوان مثال، ممکن است آدرس دانشجوی تغییر کند و این تغییر در داده‌های امور آموزشی لحاظ شود اما در بقیه قسمت‌های سیستم بدون تغییر باقی بماند.**

- مسئله دسترسی به اطلاعات:** فرض کنید یکی از کارشناسان آموزشی دانشگاه می‌خواهد لیست تمام دانشجویانی را که در یک ترم خاص معدل بالای ۱۸ دارند و همزمان در رشته‌های خاصی تحصیل می‌کنند، تهیه کند. او از بخش IT درخواست می‌کند تا چنین لیستی را تولید کنند. با این حال، سیستم فعلی دانشگاه انتظار چنین تقاضایی را نداشته است و هیچ برنامه کاربردی برای تأمین این نیاز ایجاد نشده است. اگرچه یک برنامه کاربردی وجود دارد که لیست کامل دانشجویان را بر اساس معدل یا رشته تحصیلی ارائه می‌دهد، اما کارشناس آموزشی دو انتخاب نامناسب دارد:

- **دستکاری دستی:** او باید لیست کامل دانشجویان را بگیرد و سپس دستی فیلتر کند تا دانشجویان مورد نظر خود را پیدا کند.
- **درخواست برنامه نویسی جدید:** او باید از برنامه نویس بخواهد که یک برنامه جدید بنویسد تا بتواند این لیست خاص را تولید کند.
- تکرار مشکل:

- چند روز بعد همین کارشناس بخواهد لیست دانشجویانی که حداقل دو بار دروس تکراری گذرانده‌اند و ساکن خوابگاه هستند، تهیه کند.
- باز هم هیچ برنامه‌ای برای این کار وجود ندارد و دو گزینه نامناسب پیش روی او قرار دارد: یا خود دستی انجام دهد یا دوباره از کسی بخواهد برنامه بنویسد.



- مسئله امنیت:** در سیستم‌های فایلینگ، کنترل دسترسی به داده‌ها بسیار دشوار است، در حالیکه تمامی کاربران سیستم پایگاه داده نباید قادر به دسترسی به تمامی داده‌ها باشند. به عنوان مثال، در یک دانشگاه، کارکنان بخش آموزش فقط نیاز دارند تا به بخشی از داده‌ها دسترسی داشته باشند.

- عدم وجود یک سیستم کنترل متمرکز روی کل داده‌های سازمان:** عدم وجود یک سیستم کنترل متمرکز بر روی کل داده‌ها باعث می‌شود که مدیریت و نظارت بر اطلاعات به صورت منسجم و هماهنگ انجام نشود. با نبود سیستم متمرکز، تصمیمات ممکن است بر اساس اطلاعات ناقص یا ناسازگار گرفته شود.

این مشکلات سبب آن شد تا سیستم‌های پایگاه داده توسعه یابند.



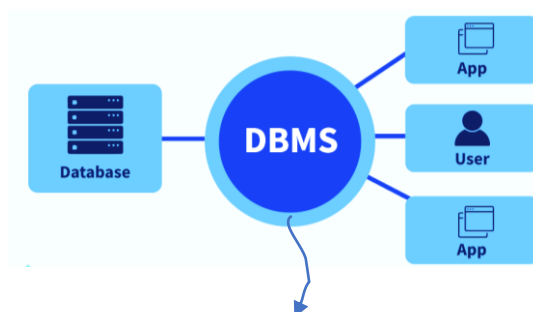


در رویکرد پایگاهی، کلیه داده‌ها به صورت یکپارچه و منسجم در یک پایگاه داده «بانک داده» ذخیره می‌شوند. این روش امکان دسترسی مستقیم به داده‌ها را برای کاربران مختلف فراهم می‌کند، در حالی که هر کاربر **دید** خاص خود را نسبت به همان مجموعه داده دارد. با استفاده از این رویکرد، افزونگی داده‌ها به حداقل ممکن کاهش می‌یابد.

نرم افزار مدیریت پایگاه داده «DBMS» نقش واسطه‌ای بین برنامه‌های کاربردی و پایگاه داده ایفا می‌کند. DBMS امکانات لازم برای مدیریت، سازماندهی و دسترسی داده‌ها را فراهم کرده و تضمین می‌کند که کاربران بتوانند به طور موثر با پایگاه داده کار کنند.



استفاده از DBMS تضمین‌کننده انسجام و سازگاری داده‌ها است؛ بدین معنی که تغییرات اعمال شده توسط هر کاربر بدون تداخل با سایر کاربران انجام خواهد شد.



واسط بین برنامه‌های کاربردی و کاربران با پایگاه داده

سیستم مدیریت پایگاه داده «database management system»

یک سیستم نرم افزاری برای ایجاد و مدیریت پایگاه‌های داده است. یک DBMS کاربران نهایی را قادر می‌سازد تا داده‌ها را در یک پایگاه داده ایجاد کنند، بخوانند، به‌روزرسانی و حذف کنند.

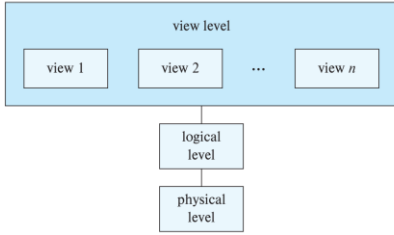


- کاهش افزونگی داده‌ها: با استفاده از یک پایگاه داده واحد، تکرار داده به حداقل می‌رسد و این باعث صرفه‌جویی در منابع و افزایش دسترسی به داده‌ها می‌شود.
- امکان اشتراک‌گذاری داده‌ها: کاربران مختلف می‌توانند به طور هم‌زمان از یک پایگاه داده استفاده کنند، که این امر همکاری را تسهیل می‌کند.
- دسترسی آسان و سریع: دسترسی به داده‌ها با استفاده از زبان‌هایی مانند SQL آسان‌تر شده است، که این امر یافتن اطلاعات مورد نظر را تسهیل می‌کند.
- امنیت بالا: امنیت داده‌ها در مقابل دسترسی غیرمجاز افزایش یافته است؛ همچنین امکان تنظیم سطوح مختلف دسترسی برای کاربران وجود دارد.



# انتزاع داده‌ها و معماری پایگاه داده

سیستم پایگاه داده از معماری خاصی برخوردار است که نحوه انتزاع داده‌ها در سطوح مختلف را مشخص می‌کند.



سطوح معماری

- **سطح فیزیکی «داخلی»:** مربوط به ذخیره‌سازی فیزیکی داده‌ها است.
- **سطح منطقی «ادراکی»:** بیانگر ساختار کلی پایگاه داده و روابط بین داده‌هاست.
- **سطح دید «خارجی»:** دیدهای کاربران نسبت به بخشی از پایگاه داده‌هاست که بر اساس نیازهایشان تعریف می‌شود.

سطح دید

- میزان دسترسی کاربران به پایگاه داده مشخص و تعریف می‌شود «به ازای هر کاربر یک دید».
- هر کاربر فقط به میزان داده‌هایی که برای انجام وظایف خود نیاز دارد، دسترسی داشته باشد.
- تنها سطحی است که به کاربران مربوط است.
- بالاترین سطح انتزاع در معماری پایگاه داده

**قانون پنهان سازی اطلاعات در اینجا صادق است: "به هر کسی به همون اندازه که اطلاعات نیاز دارد بده، نه بیشتر"**



مثال: محیط عملیاتی بانک

کاربران مختلف:

- کارمند شعبه: فقط باید به اطلاعات مربوط به مشتریان شعبه خود دسترسی داشته باشد.
- مدیر شعبه: علاوه بر اطلاعات مشتریان شعبه خود، باید بتواند گزارشات کلی از عملکرد تمام کارمندان را مشاهده کند.
- مدیر کل بانک: باید بتواند تمام داده‌های بانکی شامل گزارشات مالی و عملکرد همه شعب را ببیند.

سطح دسترسی:

- سطح دسترسی کارمند محدود است و فقط شامل دیدن حساب‌های مشتری در شعبه اوست.
- سطح دسترسی مدیر شعبه گسترده‌تر است و شامل دیدن عملکرد کارمندان نیز می‌شود.
- سطح دسترسی مدیر کل کامل‌ترین سطح است که امکان نظارت بر همه شعب را فراهم می‌کند.

مدل رابطه‌ای

- معروف‌ترین و موفق‌ترین مدل پایگاه داده است. اکثر سیستم‌های پایگاه داده کنونی بر روی مدل رابطه‌ای بنا شده‌اند.
- از مجموعه‌ای از جداول برای نمایش داده‌ها و ارتباط بین آنها استفاده می‌کند.
  - داده‌ها در قالب جدول‌ها و روابط بین جدول‌ها در چارچوب جبر رابطه‌ای سازمان‌دهی می‌شوند.
- به یک جدول یک رابطه نیز می‌گویند.
- نمونه‌های مطرح از DBMS‌های مبتنی بر مدل رابطه‌ای:
  - PostgreSQL - SQLite - Oracle - MySQL - MariaDB - Microsoft SQL Server



پایگاه داده‌ای رابطه‌ای به پایگاه داده‌ای گفته می‌شود که مبتنی بر مدل رابطه‌ای باشد.

پایگاه داده‌ای که لایه‌ی منطقی «ادراکی» آن بر اساس مدل رابطه‌ای طراحی شده باشد.

در یک جدول، داده‌ها به صورت سطرها و ستون‌هایی سازماندهی می‌شوند. هر سطر که یک واحد کامل از داده‌ها را در بر دارد، به عنوان یک **رکورد** شناخته می‌شود. این رکوردها از واحدهای کوچک‌تری تشکیل شده‌اند که به نام **فیلد** شناخته می‌شوند. در واقع، هر ستون نمایانگر یک فیلد خاص است و هر رکورد شامل چندین فیلد است.

فیلد

نام	رشته	محل تولد
سامان	علوم کامپیوتر	تهران
ساسان	مهندسی کامپیوتر	مشهد
اشکان	ریاضی	کرج
داریوش	آمار	شیراز

## طرح‌واره (سِما) پایگاه داده

به مجموعه ساختارهای طراحی شده در یک پایگاه داده بدون توجه به داده‌هایی که در آن قرار می‌گیرند، طرح‌واره «Schema» پایگاه داده گویند. به عنوان مثال در مدل رابطه‌ای که فقط با رابطه «جدول» سروکار دارد، طرح‌واره یک پایگاه داده را جدول، ستون‌ها و ارتباط بین آنها را تشکیل می‌دهند. هر یک از مقادیری که به فیلدها نسبت داده می‌شود **نمونه** گفته می‌شود.

طرح‌واره

نمونه	شماره دانشجویی
۱۸	۱۰۰۲
۱۹	۱۰۰۴

نمونه

نوع داده هر ستون «عدد»، «رشته»، ... به طرح‌واره پایگاه داده مرتبط است، ولی تعداد سطرها موجود در جدول ربطی به طرح‌واره پایگاه داده ندارد.



**ترانزاکشن**

- هر برنامه‌ای که توسط کاربران در پایگاه داده اجرا شود یک **ترانزاکشن** است. مانند عملیات کارت به کارت در یک دستگاه خودپرداز.
- عموماً ترانزاکشن‌ها دنباله‌ای از عملیات بر پایگاه داده هستند که روی هم رفته انجام یک کار یا یک وظیفه را برعهده دارند.
- یک ترانزاکشن همواره توسط DBMS کنترل می‌شود.
- تمامی ترانزاکشن‌ها یا به صورت موفق «commit» یا ناموفق «abort» پایان می‌یابند.

هر ترانزاکشنی که در سیستم اجرا می‌شود باید دارای چهار ویژگی باشد (معروف به ACID):

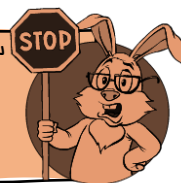
۱. یکپارچگی (Atomicity)
۲. سازگاری (Consistency)
۳. انزوا (Isolation)
۴. ماندگاری (Durability)

اولین ویژگی یک ترانزس این است که اثری که بر پایگاه داده می‌گذارد، کامل و بدون نقص باشد. به این معنی که اگر قرار است مجموعه‌ای از عملیات تغییراتی را اعمال کنند، تمام آن تغییرات باید روی جدول‌ها اجرا شوند. اگر حتی یکی از عملیات با مشکل مواجه شود، تأثیرات عملیات قبلی باید لغو شوند. به بیان ساده‌تر، در یک ترانزس یا همه عملیات باید به طور کامل انجام شوند یا هیچ کدام نباید اجرا شده و اثرگذار باشند. این ویژگی به همه یا هیچ موسوم است.

## سازگاری

ترانزس باید تغییرات را به گونه‌ای اعمال کند که پایگاه داده از یک وضعیت صحیح به وضعیت صحیح دیگری منتقل شود. این تضمین توسط قوانین جامع پایگاه داده انجام می‌شود، بنابراین ترانزس باید تغییرات را به نحوی اعمال کند که این قوانین نقض نشوند. هر ترانزسی که محدودیت‌ها را نقض کند، هرگز موفق نخواهد شد.

آیا دو خاصیت یکپارچگی و سازگاری به تنهایی برای اطمینان از صحت ترانزس‌ها کافی نیستند؟ خیر ☹️  
ممکن است ترانزس‌های هم‌روند روی یک دیگر اثر مخرب داشته باشند 😱



## اتروا جدایی

هر ترانزس باید به طور مستقل از سایر ترانزس‌هایی که ممکن است هم‌زمان در حال اجرا باشند، عمل کند. اگر خاصیت اتروا در ترانزس‌ها رعایت نشود، یعنی اگر یک ترانزس، نتایج نهایی نشده‌ی یک ترانزس دیگر را بخواند، مقدار داده‌ای را می‌خواند که هنوز سازگار نشده است. در نتیجه، اجرای ترانزس دوم به نتیجه‌ای ناسازگار منجر خواهد شد. براساس این ویژگی، هم‌روندی ترانزس‌ها باید کنترل شود تا اثر مخرب بر روی هم نداشته باشند.

**مثال:** دو ترانزس هم‌زمان در حال انجام هستند؛ یکی سفارش یک محصول توسط مشتری A و دیگری سفارش همان محصول توسط مشتری B.

**مشکل بدون اتروا:** اگر خاصیت اتروا رعایت نشود، ممکن است سیستم قبل از اینکه موجودی کالا به روز شود، سفارش دوم را تأیید کند و منجر به فروش بیس از موجودی شود.

**راه حل با اتروا:** با استفاده از خاصیت اتروا، ترانزس‌ها به صورت مستقل اجرا می‌شوند. یعنی اگر مشتری اول سفارش دهد، سیستم موجودی را چک کند و تا زمانی که ترانزس اول تمام نشده باشد، مثلاً تا زمان پرداخت نهایی، دیگر مشتریان نمی‌توانند آن کالا را بخرند. پس از اتمام ترانزس اول و به روز شدن موجودی، سپس مشتری دوم می‌تواند اقدام به خرید کند.

## روش قفل‌گذاری

هنگامی که یک ترانزس به داده‌ای نیاز داشته باشد، تقاضای قفل کردن آن را می‌دهد و در این وضعیت باقی ترانزس‌ها تا اتمام کار آن ترانزس نمی‌توانند از آن استفاده کنند.

خاصیت ماندگاری یک ترانسس یعنی نتایج اجرای موفق یک ترانسس باید اثرش ماندنی و پایدار باشد، حتی اگر پس از اجرای ترانسس نقصی در سیستم بروز کند نتایج ترانسس قبلی از بین نمی رود و قابل دسترس باقی می ماند. به عنوان مثال، فرض کنید شما مبلغی را از حساب خود به حساب شخص دیگری انتقال می دهید. اگر این ترانسس با موفقیت انجام شود و سپس سیستم دچار مشکل شود یا خاموش گردد، خاصیت ماندگاری تضمین می کند که انتقال وجه بعد از راه اندازی مجدد سیستم نیز معتبر باشد و مبلغ منتقل شده در حساب مقصد ثبت شده باقی بماند یا در مثالی دیگر، فرض کنید که شما یک سفارش آنلاین می دهید و پرداخت آن را با کارت بانکی انجام می دهید. پس از تأیید پرداخت، سیستم فروشگاه آنلاین سفارش شما را ثبت می کند. اگر درست بعد از این که سیستم فروشگاه آنلاین سفارشات را ثبت کرد، به دلیل مشکل فنی یا قطع برق خاموش شود، خاصیت ماندگاری تضمین می کند که اطلاعات سفارش و پرداخت شما در سیستم ثبت شده باقی بماند.

## طراحی پایگاه داده

فرآیند تولید یک سیستم پایگاهی در یک محیط عملیاتی، شامل چند مرحله مهم است که به طور کلی می توان آن را به چهار گام اصلی تقسیم کرد: نیازسنجی، مهندسی خواسته ها، مدل سازی داده، طراحی و پیاده سازی. در یک نرم افزار سیستم مدیریت پایگاه داده.

۱. نیازسنجی: در این مرحله، طراحان باید با کاربران و ذی نفعان مختلف مشورت کنند تا نیازهای سیستم را به درستی درک کنند که

شامل جمع آوری داده های مورد نظر برای ذخیره سازی و تعیین چگونگی استفاده کاربران از داده ها است. (این مرحله در درس مهندسی نرم افزار مطرح می گردد)

▪ تعیین اهداف: مشخص کردن هدف اصلی پایگاه داده.

▪ جمع آوری داده ها: شناسایی انواع داده هایی که باید ذخیره شوند.

▪ تحلیل انتظارات: درک چگونگی دسترسی کاربران به داده ها.

۲. مدل سازی داده ها: در این مرحله، نیازهای داده ای جمع آوری شده در مرحله قبل به یک مدل مفهومی تبدیل می شود. این مدل مفهومی

به صورت نمودار موجودیت-رابطه نمایش داده می شود و موجودیت ها و روابط بین آن ها را به صورت گرافیکی نشان می دهد.

▪ تعیین موجودیت ها: شناسایی موجودیت های اصلی و ویژگی های آنها

▪ تعیین روابط: شناسایی روابط بین موجودیت ها

▪ تعیین محدودیت ها: اعمال محدودیت های داده ای مانند یکتایی، التزامی بودن و ...

۳. طراحی: در این مرحله، مدل مفهومی به یک مدل منطقی تبدیل می شود که قابل پیاده سازی در یک سیستم مدیریت پایگاه داده

DBMS است. این مرحله شامل تعیین ساختار جدول ها، ستون ها، انواع داده ها، و روابط بین جدول ها است.

▪ تبدیل موجودیت ها به جدول ها: هر موجودیت در مدل مفهومی به یک جدول در مدل منطقی تبدیل می شود.

▪ تبدیل ویژگی ها به ستون ها: هر ویژگی از موجودیت ها به یک ستون در جدول تبدیل می شود.

▪ تعیین روابط بین جدول ها: روابط بین موجودیت ها به روابط بین جدول ها تبدیل می شوند و کلیدهای خارجی تعیین می شوند.

۴. پیاده سازی: در این مرحله، طراحی منطقی پایگاه داده به یک پایگاه داده فیزیکی تبدیل می شود که قابل استفاده توسط برنامه های

کاربردی است. این مرحله شامل ایجاد جدول ها، تعیین محدودیت ها، و وارد کردن داده های اولیه است.