



d-learn.ir

درس SQL برای تحلیل داده

فصل 1: مبانی پایگاه‌های داده

مجید پورکاشانی



دقیقه

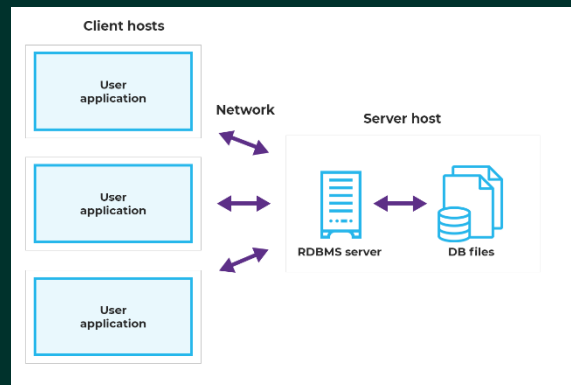
مدرسه پردازش و تحلیل داده

موضوعات جلسه امروز

- آشنایی با پایگاه‌های داده
 - پایگاه داده چیست؟
 - دیتابیس‌های رابطه‌ای (Relational Database)
 - انواع RDBMS و پایگاه داده‌های خاص
 - زبان SQL
 - تاریخچه سیکوئل سرور
 - ویرایش‌های سیکوئل سرور
 - اجزای سیکوئل سرور
 - ابزارها
- ساختار دیتابیس (بانک اطلاعاتی)

پایگاه داده

- سیستم مدیریت پایگاه داده (Database Management System) یک سیستم نرم‌افزاری برای ایجاد، نگهداشت و مدیریت داده‌هاست که از اجرای مختلف تشکیل شده و معمولاً روی یک سرور نصب و راه‌اندازی می‌شود. DBMS تعداد دلخواهی از پایگاه‌های داده (Database) را در خود نگهداری می‌کند.
- هر پایگاه داده (Database) مجموعه‌ای متشکل و سازگار از داده‌هاست که بر مبنای نوعی از محدودیت‌ها امکان ذخیره‌سازی و دسترسی ساختاریافته داده‌ها را می‌دهد.



پایگاه داده

- ویژگی‌های پایگاه داده
 - دسترسی همزمان چندین کاربر
 - امنیت و سطوح دسترسی
 - ذخیره‌سازی مطمئن
 - حفظ سازگاری و تمامیت داده‌ها
- سیستم مدیریت پایگاه داده عموماً از تکنولوژی کلاینت-سرور استفاده می‌کند و روی شبکه قابل دسترسی است.

کاربردهای پایگاه داده

- بانکداری: مدیریت حساب‌های مشتریان، تراکنش‌ها و اطلاعات مالی.
- تجارت الکترونیک: مدیریت محصولات، سفارشات مشتریان و اطلاعات پرداخت.
- سلامت: مدیریت سوابق پزشکی بیماران، داروها و آزمایش‌ها.
- حمل و نقل: مدیریت برنامه‌های پرواز، رزرواسیون و اطلاعات مسافران.
- تولید: مدیریت موجودی کالاها، فرآیند تولید و کنترل کیفیت.

مزایا استفاده از پایگاه داده



• یکپارچگی داده‌ها

- حذف داده‌های تکراری و تضمین سازگاری اطلاعات در کل سیستم

• امنیت داده‌ها

- کنترل دسترسی به داده‌ها و جلوگیری از دسترسی‌های غیرمجاز.
- امکان پشتیبان‌گیری و بازیابی داده‌ها در صورت بروز مشکل.

• دسترسی سریع و آسان

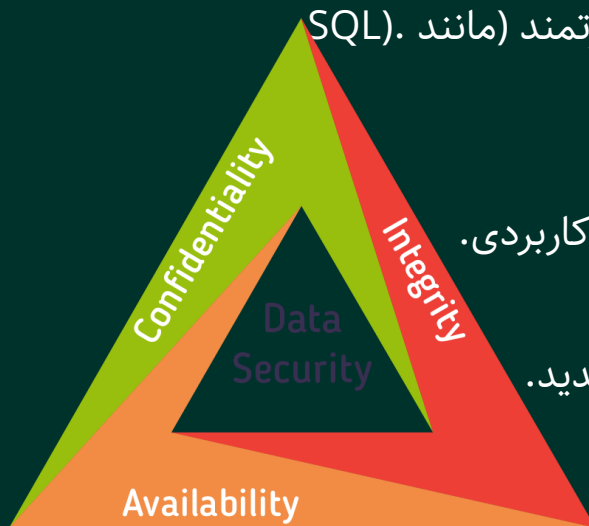
- جستجوی سریع اطلاعات با استفاده از زبان‌های پرس‌وجوی قدرتمند (مانند SQL).
- امکان ایجاد گزارش‌ها و تحلیل‌های مختلف بر روی داده‌ها.

• اشتراک‌گذاری داده‌ها

- امکان اشتراک‌گذاری اطلاعات بین کاربران مختلف و برنامه‌های کاربردی.

• انعطاف‌پذیری

- قابلیت تغییر و توسعه پایگاه داده برای پاسخگویی به نیازهای جدید.



پایگاه داده رابطه‌ای

- پایگاه داده رابطه‌ای (Relational Database)
- معرفی شده توسط E.F. Codd (کارمند IBM) در سال 1970
- دیتا با مجموعه‌ای از چندتایی‌ها (Tuple) های $(d_1, d_2, \dots, d_n)$ مشخص می‌شود.
- هر d_i عضوی از یک قلمرو (Domain) داده ای D_i است.

نمایش دیتا به فرم رابطه

Tuple	Attribute				
	D1	D2	D3	D4	D5
	...	...	...	...	...
	d1	d2	d3	d4	D5
	...	...	...	...	...
Relation					

برخی از مهمترین RDBMS ها



Oracle -

MSSQL Server -

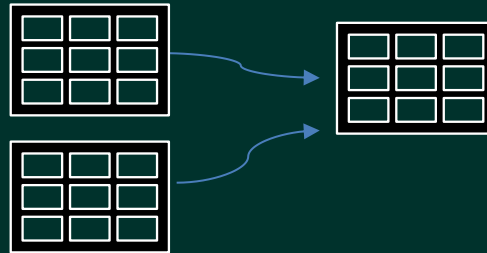
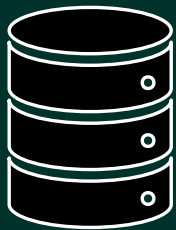
MySQL / MariaDB -

Postgres -

SQLite -

IBM DB2 -

ساختار پایگاه داده



- پایگاه داده

- جدول

- رکورد

- فیلد

ID	Name	Age	Gender	Salary	Phone	Address
1	John Doe	34	M	125,000	+1313421 ...	#1 Black St.
2	Jane Doe	45	F	140,000		#7 Brown St
3	...	...	...	...	...	...
4	...	...	...	...	...	...
5	...	...	...	...	...	...

ساختار پایگاه داده

- رکورد

- معادل یک جزء مشاهده (Observation)، تراکنش (Transaction)، اتفاق (Event) یا وضعیت (State) در دنیای واقعی
- دارای ویژگی (Attribute) های مشخص
- مثال: اطلاعات یک دانش آموز (Student)

ID	Name	Age	BirthDate
132	Ali Saberi	35	1378/12/03

- فیلد

- کوچکترین عنصر اطلاعاتی که جزو اجزای رکورد است.
- معادل ویژگی (Attribute) یک پدیده، مشاهده، تراکنش
- دارای یک Type یا نوع ثابت
- متنی، عددی، جغرافیایی، باینری، تاریخ
- مثال: سن دانش آموز (Age)

نوع داده (Data Type)

- بر اساس تعریف پایگاه داده رابطه‌ای، هر فیلد (ستون) یک نوع داده مشخص و غیرقابل تغییر دارد.
- انواع داده.
 - متنی یا رشته‌ای (VARCHAR ، CHAR)
 - متنی بین‌المللی (NVARCHAR ، NCHAR)
 - باینری (BLOB ، VARBINARY ، BINARY)
 - عددی
 - دقیق (BIGINT ، INTEGER ، SMALLINT ، DECIMAL ، NUMERIC)
 - تقریبی (REAL ، FLOAT)
 - تاریخ/ساعت (TIMESTAMP ، TIME ، DATE)
 - بولی (BOOLEAN)
 - انواع پیچیده‌تر (JSON ، XML)

پایگاه داده‌های غیررابطه‌ای (NoSQL)



- گراف دیتابیس: Neo4j



Prometheus

- سری زمانی: Prometheus



redis

- کلید/مقداری: Redis



mongoDB®

- غیرساخت یافته: MongoDB



elasticsearch

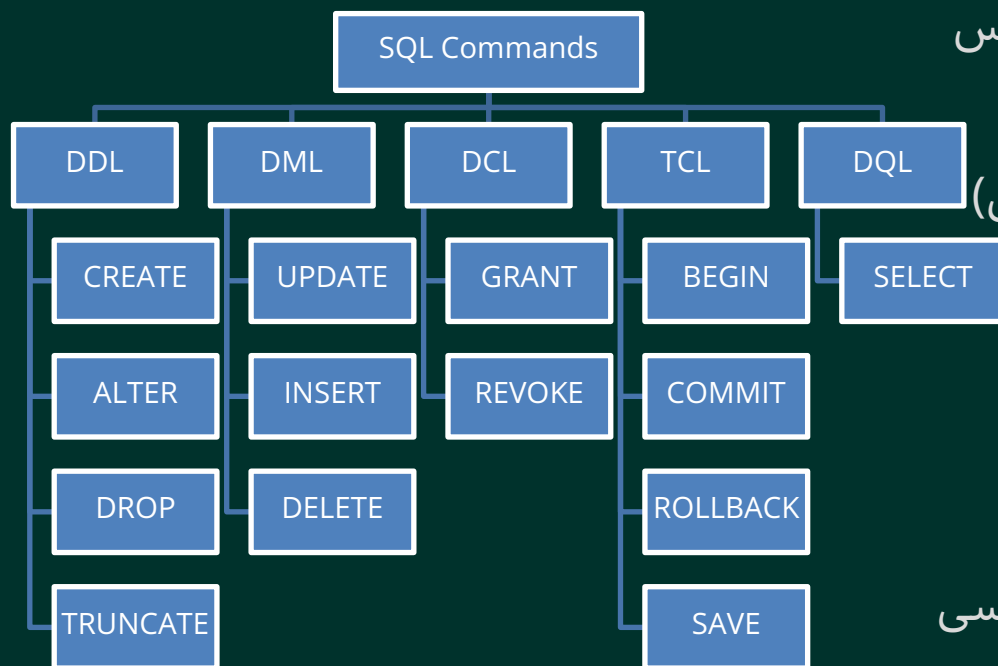
- اندیس متنی: Elasticsearch

زبان رابط برای دسترسی و تغییر دیتا

- زبان معمول برای ارتباط با RDBMS، اسکیوال است: Simple Query Language یا SQL.
- به اسکریپت‌های اجرای دستورات SQL، **کوئری** می‌گوییم.
- اسکیوال یک زبان است با **گویش**‌های مختلف

DMBS	لهجه
Oracle	PL-SQL
MS SQL Server, Sybase	T-SQL
PostgreSQL	pgSQL
mySQL	mySQL dialect
Google BigQuery	ANSI / Standard

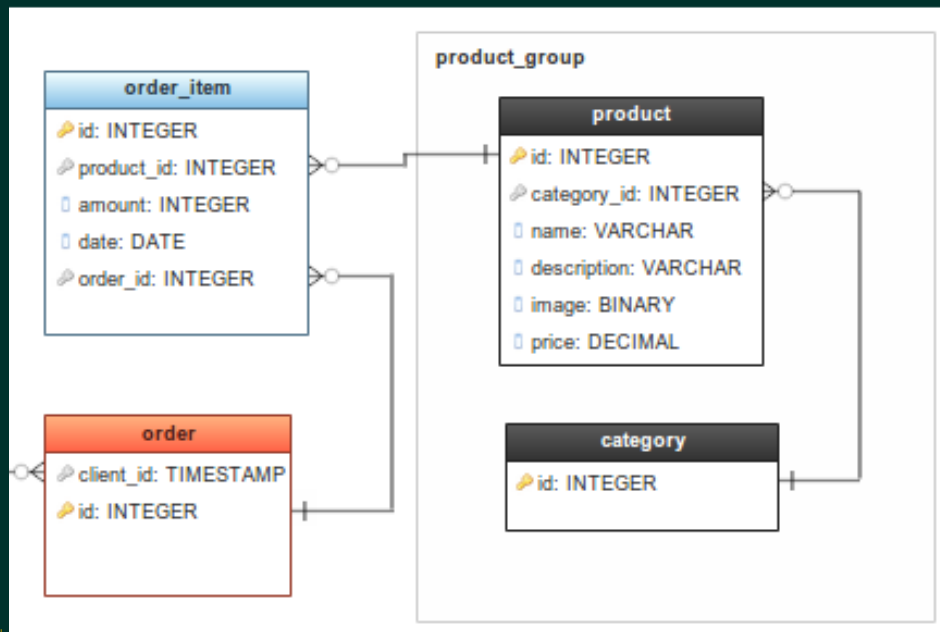
بخشهای مختلف زبان SQL



- زبان توصیف دیتا (DDL)
- دستورات ایجاد ساختار دیتابیس
- زبان بازخوانی دیتا (DQL)
- دستورات بازخوانی دیتا (کوئری)
- زبان تغییر دیتا (DML)
- دستورات تغییر، به روزرسانی، حذف و درج دیتا
- زبان کنترل دیتا (DCL)
- دستورات تعیین سطوح دسترسی
- زبان کنترل تراکنش (TCL)
- دستورات کنترل اجرای عملیات داخلی دیتابیس

ساختار پایگاه داده

- جدول‌ها معادل ماهیت (Entity) های واقعی در کسب و کار یا گردش کار سازمانی هستند.
- جداول در دیتابیس از طریق Relation با هم ارتباط دارند.
- Foreign Key / Primary Key



Database
Diagram



پایان فصل

پایگاه‌های داده رابطه‌ای

