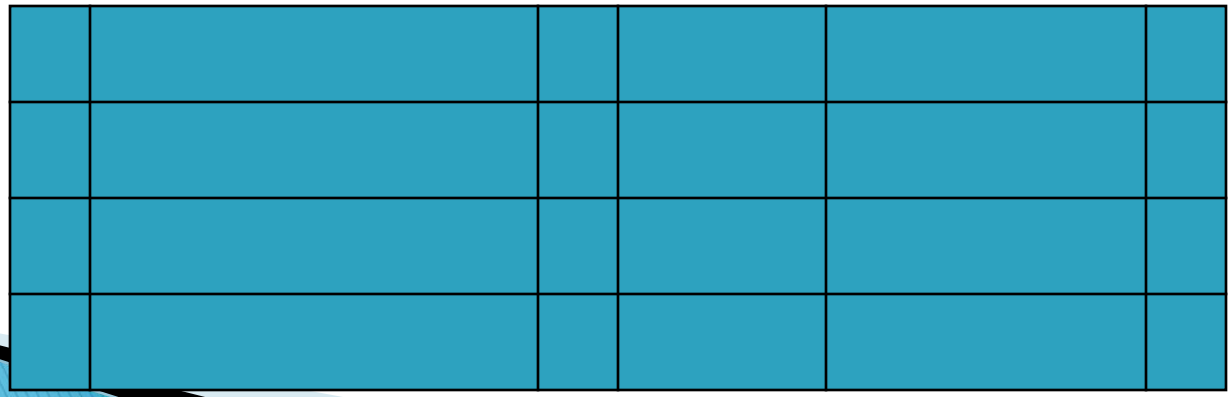
Database Tables and Their Components

Database Table Contents

- A database table is used to store an entity set
  - An entity set is a collection of related entities
    - **An entity is anything you want to keep track of, so an entity may be a person, place, thing, event, etc.**

Database Table Components

- At the *conceptual* level, a database table may be viewed as a matrix.
  - Matrix rows are also known as **tuples** or **records**
    - **each row contains an entity**
  - Matrix columns are also called **fields**.
    - **Each column (field) contains the entity's attribute values**



Keys

- Primary Key (PK)
  - an attribute (or combination of attributes) that uniquely identifies each row (entity) in a table.
    - **A PK composed of two or more attributes is known as a *composite* PK.**
- Foreign Key (FK)
  - an attribute in one table whose values match the PK values in a related table or whose “values” are null.
    - **FKs are used link (connect) related tables.**

Keys, cont.

- Entity Integrity
  - PK uniquely identifies each entity in a table
    - **PK may not include nulls**
- Referential Integrity
  - FK values in one table match the PK values in the related table
    - **FK may include nulls**

A Conceptual View of a Database Table

Tables are named. The table you see here is the EMPLOYEE table.

Each column contains the values of an attribute. The EMP_FNAME column only contains employee first names; the EMP_PHONE may only contain employee phone numbers.

Named **attributes** (**fields**)

	EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_INITIAL	EMP_ADDRESS	EMP_PHONE
▶	504	Graham	Georgette	E	1234 Trident Lane, Eaton, MA 12345	234-234-6789
	506	Smith	Alex		4321 Hill Drive, Dalton, GA 23456	123-456-7890
	507	Thieu	Robert	E	5678 Crescent Lane, Dalton, GA 23456	123-457-1123
	510	Chen	Alice	L	3456 Lake Drive, Nashville, TN 37654	456-333-9876
	512	Alazar	George	D	2345 Oak Drive, Nashville, TN 33123	456-456-7654
*	0					

Each row represents an entity

A row is also called a record or a tuple.

Each row/column intersection contains *only one* of an entity's attribute values

The Primary Key (PK) is a unique entity identifier. If you know the PK value, you will know all of its row's attribute values

Entity Integrity

EMPLOYEE table

	EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_INITIAL	EMP_ADDRESS	EMP_PHONE
▶	504	Graham	Georgette	E	1234 Trident Lane, Eaton, MA 12345	234-234-6789
	506	Smith	Alex		4321 Hill Drive, Dalton, GA 23456	123-456-7890
	507	Thieu	Robert	E	5678 Crescent Lane, Dalton, GA 23456	123-457-1123
	510	Chen	Alice	L	3456 Lake Drive, Nashville, TN 37654	456-333-9876
	512	Alazar	George	D	2345 Oak Drive, Nashville, TN 33123	456-456-7654
*	0					

A table exhibits entity integrity when all of its Primary Key (PK) values *uniquely* identify each table row (record.)

1. **A PK cannot contain duplicate values**
2. **A PK cannot contains nulls**

Note: A null indicates the ***absence*** of a value; it is ***not*** a blank. (You create a null when you tap the ENTER key without first making an entry. A blank is created when you tap the space bar and then tap the ENTER key.)

Foreign keys (FK) and Referential Integrity

A Foreign Key (FK) is an attribute located in one table that "points to" a Primary Key (PK) in a related table. The use of FKs allows you relate one table to another.

To maintain referential integrity, a foreign key (FK) must reference an *existing* PK value in a *related* table or it may be null.

INVOICE table

	INV_NUM	INV_DATE	CUS_NUM	INV_SUBTOTAL	INV_TAX	INV_TOTAL	INV_PAYMENT	INV_BALANCE
▶	100541	Friday, June 11, 1999	112	$86.90	$6.95	$93.85	$93.85	$0.00
	100542	Friday, June 11, 1999	122	$24.95	$2.00	$26.95	$26.95	$0.00
	100543	Saturday, June 12, 1999	112	$91.15	$7.29	$98.44	$98.44	$0.00
*	0		0	$0.00	$0.00	$0.00	$0.00	$0.00

LINE table

	INV_NUM	LINE_NUM	PROD_CODE	LINE_UNITS	LINE_PRICE	LINE_AMOUNT
▶	100541	1	SME-123456	8	$10.05	$80.40
	100541	2	FRT-765432	1	$6.50	$6.50
	100542	1	DFR-345678	1	$24.95	$24.95
	100543	1	SME-123456	4	$10.05	$40.20
	100543	2	DFR-345678	1	$24.95	$24.95
	100543	3	GRD-998877	4	$6.50	$26.00
*	0	0		0	$0.00	$0.00

Attribute Storage

- Each table row/column intersection contains a single attribute value for a single entity.
  - Ideally, attributes are **simple** and **single-valued**.
  - **Single-valued composite attributes** are acceptable
    - **but composite attributes may make queries more complex and may impose reporting limitations.**
  - **Multi-valued attributes**, either simple or composite, may …
    - **create structural problems**
    - **make queries more complex and may impose reporting limitations.**

Multi-valued Attribute Storage

- **Multi-valued attributes** are sometimes stored as
  - **strings**
    - this approach yields query complexity and reporting limitations
  - **multiple attributes**
    - this approach yields structural problems
- Ideally, multi-valued attributes are handled through the use of composite tables.

Multi-valued Attributes Stored As Strings

Multi-valued attribute

Poor practice: Makes it difficult to generate queries such as "How many employees have earned BA or MBA degrees?"

	EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_INITIAL	EMP_DEGREE
▶	458	Renselaer	Randolph	D	BA
	459	Freeman	William	H	
	460	Hartog	Susanne	M	
	462	Appleton	Anne	M	BBA, MBA, MFA
	463	Randall	Herman	A	BE
	464	Hernandez	Olanzo	R	BS, MS, MBA, Ph.D.
	465	Oszwicki	James	P	BS, MS
*	0				

Storing Multi-valued Attributes In Separate Columns

Poor structure:

- **many nulls**
- **table structure must be altered when additional degrees are earned**

	EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_INITIAL	EMP_DEGREE1	EMP_DEGREE2	EMP_DEGREE3	EMP_DEGREE4
▶	458	Renselaer	Randolph	D	BA			
	459	Freeman	William	H				
	460	Hartog	Susanne	M				
	462	Appleton	Anne	M	BBA	MBA	MFA	
	463	Randall	Herman	A	BE			
	464	Hernandez	Olanzo	R	BS	MS	MBA	Ph.D.
	465	Oszwicki	James	P	BS	MS		
*	0							

A Composite Table (EDUCATION) is Used To Convert M:N Relationships To 1:M Relationships

EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_INITIAL
458	Renselaer	Randolph	D
459	Freeman	William	H
460	Hartog	Susanne	M
462	Appleton	Anne	M
463	Randall	Herman	A
464	Hernandez	Olanzo	R
465	Oszwicki	James	P
0			

Table name: EMPLOYEE

Table name: EDUCATION

EMP_NUM	DEGREE_CODE	EDUC_EARNED	EDUC_INSTITUTION
458	BA	1993	U. of Kentucky
462	BBA	1985	U. of Massachussetts
462	MBA	1989	UCLA
462	MFA	1996	U. of Memphis
463	BA	1992	U. of Florida
464	BS	1981	Florida State University
464	MBA	1991	Penn Sate University
464	MS	1986	Penn State University
464	Ph.D.	1999	Michigan State University
465	BS	1985	Cal Tech
465	MS	1986	Cal Tech
0			

Table name: DEGREE

DEGREE_CODE	DEGREE_NAME
BA	Bachelor of Arts
BBA	Bachelor of Business Administration
BS	Bachelor of Science
MBA	Master of Business Administration
MFA	Master of Fine Arts
MS	Master of Science
Ph.D.	Doctor of Philosophy

Supertype/Subtype Relationships Are Used To Eliminate or Control the Occurrence of Nulls

	EMP_NUM	EMP_LNAME	EMP_LICENSE	EMP_MED_TYPE	EMP_MEDICAL	EMP_FLT_CHECK
▶	1205	Smith	ATP	1	Tuesday, February 16, 1999	Monday, March 22, 1999
	1211	Jones				
	1212	Hernandez	ATP	1	Friday, April 23, 1999	Wednesday, May 19, 1999
	1213	Wuizicki	COM	2	Friday, December 18, 1998	Thursday, July 15, 1999
	1216	Anderson				
*	0					

A table with many nulls

	EMP_NUM	EMP_LNAME
▶	1205	Smith
	1211	Jones
	1212	Hernandez
	1213	Wuizicki
	1216	Anderson
*	0	

Supertype. (Table name: EMPLOYEE

Subtype. (Table name: PILOT)

	EMP_NUM	EMP_LICENSE	EMP_MED_TYPE	EMP_MEDICAL	EMP_FLT_CHECK
▶	1205	ATP	1	Tuesday, February 16, 1999	Monday, March 22, 1999
	1212	ATP	1	Friday, April 23, 1999	Wednesday, May 19, 1999
	1213	COM	2	Friday, December 18, 1998	Thursday, July 15, 1999
*	0				

The END