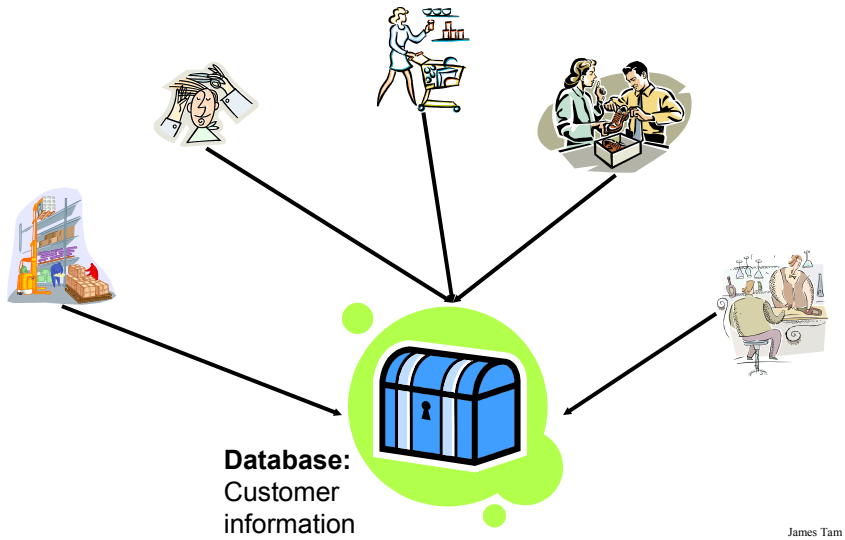
Databases

In this section of notes you will learn about: different types of databases, how information is stored in databases, the different types of relations that can exist within a database and how information can be retrieved via queries.

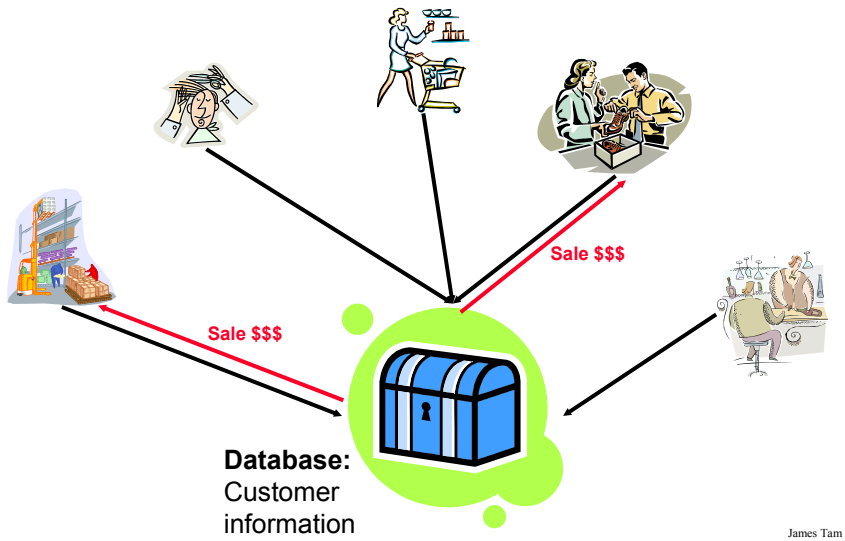
Purpose Of A Database

To store information



Purpose Of A Database

To retrieve information



Judging Databases

- In and of itself the storage of information is neither good nor bad.
- What should be judged:
  - How is the information used
  - Security of the information

Storing Information In A Database

Information is commonly stored in tables (relational database):

'Employees' table

SIN	LastName	FirstName	Address	City	Province
638666670	Cartland	Douglas	1109, 4944 Dalworth Dr	Silent Hill	Alberta
456789123	Cartman	Eric	456 Lynchview Road	Southpark	Alberta
670380456	Edgar	Maureen	300, Lockinvar Road	Calgary	Alberta
456889123	Flanders	Ned	60 Evergreen Terrace	Springfield	Alberta
413754621	Kennedy	Leon	808, 4900 Wildman Ave	Racoon City	Alberta
456438624	Lemoy	Leonard	55 Logic Way	Vulcan	Alberta
666666667	Mason	Harry	7 Luckstone Dr	Silent Hill	Alberta
666666666	Morris	Heather	7 Luckstone Dr	Silent Hill	Alberta
444638047	Redfield	Claire	653 Wildpark Place	Racoon City	Alberta
123115323	Simcox	Cole	311 Ocean View Drive	Vancouver	British C
456789124	Simpson	Homer	59 Evergreen Terrace	Springfield	Alberta
123456789	Smith	John	123 Peanut Lane	Calgary	Alberta
666666668	Sunderland	James	7 Heartbroken Ave	Silent Hill	Alberta
620451097	Williams	Amanda	25 Rodeo Drive	Edmonton	Alberta
666666669	Wolf	Claudia	66 Twisted View	Silent Hill	Alberta
371988812	Carswell	Mary	425 Remington Ave	Calgary	Alberta

Storing Information In A Database (2)

Record: An example instance of data within the table.

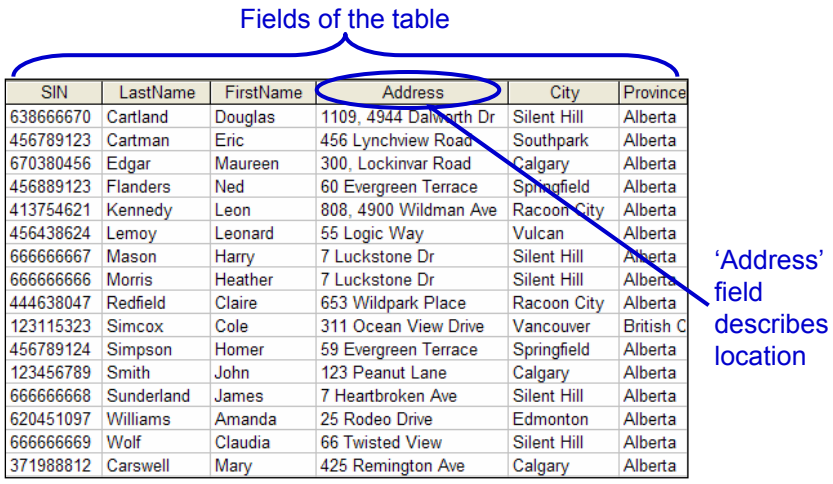
Records of the table

SIN	LastName	FirstName	Address	City	Province
638666670	Cartland	Douglas	1109, 4944 Dalworth Dr	Silent Hill	Alberta
456789123	Cartman	Eric	456 Lynchview Road	Southpark	Alberta
670380456	Edgar	Maureen	300, Lockinvar Road	Calgary	Alberta
456889123	Flanders	Ned	60 Evergreen Terrace	Springfield	Alberta
413754621	Kennedy	Leon	808, 4900 Wildman Ave	Racoon City	Alberta
456438624	Lemoy	Leonard	55 Logic Way	Vulcan	Alberta
666666667	Mason	Harry	7 Luckstone Dr	Silent Hill	Alberta
666666666	Morris	Heather	7 Luckstone Dr	Silent Hill	Alberta
444638047	Redfield	Claire	653 Wildpark Place	Racoon City	Alberta
123115323	Simcox	Cole	311 Ocean View Drive	Vancouver	British C
456789124	Simpson	Homer	59 Evergreen Terrace	Springfield	Alberta
123456789	Smith	John	123 Peanut Lane	Calgary	Alberta
666666668	Sunderland	James	7 Heartbroken Ave	Silent Hill	Alberta
620451097	Williams	Amanda	25 Rodeo Drive	Edmonton	Alberta
666666669	Wolf	Claudia	66 Twisted View	Silent Hill	Alberta
371988812	Carswell	Mary	425 Remington Ave	Calgary	Alberta

One record, 'Simpson, Homer'

Storing Information In A Database (3)

Field: are attributes used to describe each record in a table



SIN	LastName	FirstName	Address	City	Province
638666670	Cartland	Douglas	1109, 4944 Dalhurth Dr	Silent Hill	Alberta
456789123	Cartman	Eric	456 Lynchview Road	Southpark	Alberta
670380456	Edgar	Maureen	300, Lockinvar Road	Calgary	Alberta
456889123	Flanders	Ned	60 Evergreen Terrace	Springfield	Alberta
413754621	Kennedy	Leon	808, 4900 Wildman Ave	Racoon City	Alberta
456438624	Lemoy	Leonard	55 Logic Way	Vulcan	Alberta
666666667	Mason	Harry	7 Luckstone Dr	Silent Hill	Alberta
666666666	Morris	Heather	7 Luckstone Dr	Silent Hill	Alberta
444638047	Redfield	Claire	653 Wildpark Place	Racoon City	Alberta
123115323	Simcox	Cole	311 Ocean View Drive	Vancouver	British C
456789124	Simpson	Homer	59 Evergreen Terrace	Springfield	Alberta
123456789	Smith	John	123 Peanut Lane	Calgary	Alberta
666666668	Sunderland	James	7 Heartbroken Ave	Silent Hill	Alberta
620451097	Williams	Amanda	25 Rodeo Drive	Edmonton	Alberta
666666669	Wolf	Claudia	66 Twisted View	Silent Hill	Alberta
371988812	Carswell	Mary	425 Remington Ave	Calgary	Alberta

Tables For Your Second Assignment

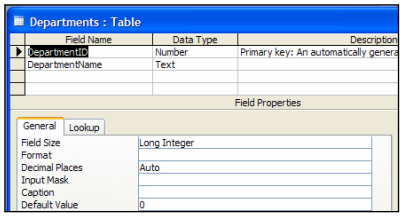
Employees table
- SIN
- LastName
- FirstName
- Address
- City
- Province
- PostalCode
- HomePhone
- BirthDate
- PayRate

Departments table
- DepartmentID
- DepartmentName

TimeBilled table
- TimeBilledID
- EmployeeID
- DepartmentID
- StartPayPeriod
- HoursWorked

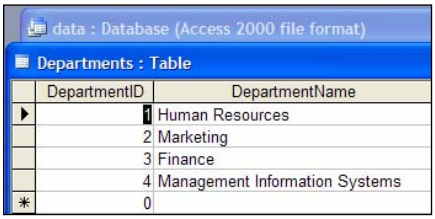
MS-Access: Views Of Your Database

Design view



- Typically start with this view
- Used to specify what fields that a table will consist of:
  - e.g., DepartmentID, DepartmentName
- Used to specify the type and the format of the information in each field:
  - e.g., SIN is field with 9 characters that must be in the format 000 000 000

Datasheet view



- Once the fields have been specified in the Design view using the Datasheet view allows for each record to be entered.

Types Of Tables

Data tables

- Stores data that provides information about the database
- Dynamic, will likely be manipulated over the life the database (add, delete, modify)
- E.g. for A2, Employees, TimeBilled tables

Validation tables

- Used to ensure data integrity (to 'lookup' values)
- Typically it maps one value to another (e.g., product to product code, book to ISBN number)
- Rarely (if ever) changes
- E.g., for A2 Departments table

DepartmentID	DepartmentName
1	Human Resources
2	Marketing
3	Finance
4	Management Information Systems

Primary Key

Each table should typically have one field designated as the primary key:

- The primary key must be unique
- It uniquely identifies one record from another

SIN	LastName	FirstName	Address	City	Province
638666670	Cartland	Douglas	1109, 4944 Dalworth Dr	Silent Hill	Alberta
456789123	Cartman	Eric	456 Lynchview Road	Southpark	Alberta
670380456	Edgar	Maureen	300, Lockinvar Road	Calgary	Alberta
456889123	Flanders	Ned	60 Evergreen Terrace	Springfield	Alberta
413754621	Kennedy	Leon	808, 4900 Wildman Ave	Racoon City	Alberta
456438624	Lemoy	Leonard	55 Logic Way	Vulcan	Alberta
666666667	Mason	Harry	7 Luckstone Dr	Silent Hill	Alberta
666666666	Morris	Heather	7 Luckstone Dr	Silent Hill	Alberta
444638047	Redfield	Claire	653 Wildpark Place	Racoon City	Alberta
123115323	Simcox	Cole	311 Ocean View Drive	Vancouver	British C
456789124	Simpson	Homer	59 Evergreen Terrace	Springfield	Alberta
123456789	Smith	John	123 Peanut Lane	Calgary	Alberta
666666668	Sunderland	James	7 Heartbroken Ave	Silent Hill	Alberta
620451097	Williams	Amanda	25 Rodeo Drive	Edmonton	Alberta
666666669	Wolf	Claudia	66 Twisted View	Silent Hill	Alberta
371988812	Carswell	Mary	425 Remington Ave	Calgary	Alberta

Primary Key for table 'Employees' is the 'SIN' field

Foreign Key

A key in one table that refers to a key in another field:

- E.g. for A2: EmployeeID and DepartmentID field of the TimeBilled table

TimeBilled : Table

TimeBilledID	EmployeeID	DepartmentID	StartPa
2	123115323	1	10/1/200
3	123456789	1	10/1/200
4	371988812	1	10/1/200
5	413754621	2	10/1/200
6	444638047	2	10/1/200
7	456438624	2	10/1/200
8	456789123	2	10/1/200
9	456789124	2	10/1/200
10	456889123	2	10/1/200
11	620451097	2	10/1/200
12	620451097	1	10/8/200
13	638666670	3	10/1/200

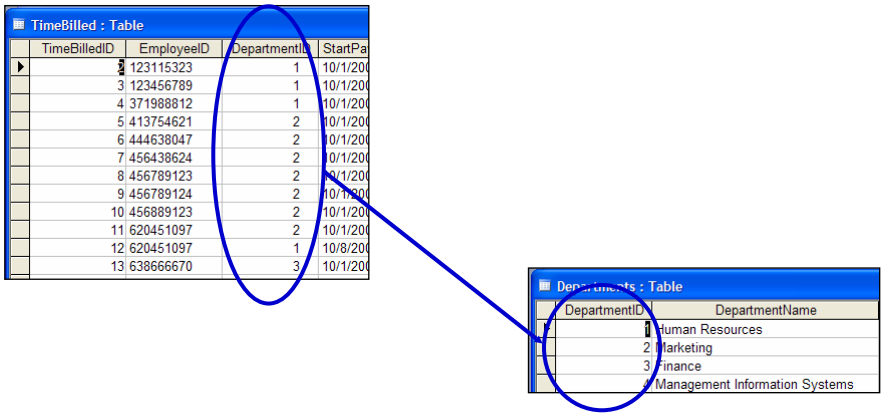
Employees : Table

SIN	LastName	FirstName	Ad
638666670	Cartland	Douglas	1109, 494
456789123	Cartman	Eric	456 Lync
670380456	Edgar	Maureen	300, Lock
456889123	Flanders	Ned	60 Evergr
413754621	Kennedy	Leon	808, 490
456438624	Lemoy	Leonard	55 Logic
666666667	Mason	Harry	7 Lucksto
666666666	Morris	Heather	7 Lucksto

Foreign Key

A key in one table that refers to a key in another field:

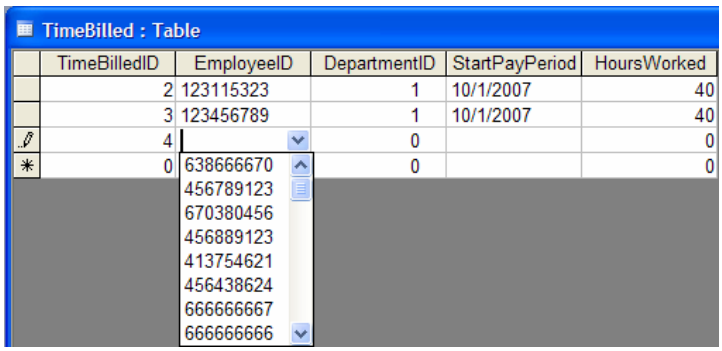
- E.g. for A2: EmployeeID and DepartmentID field of the TimeBilled table



Purpose Of Foreign Keys

- To ensure the integrity of the foreign key.

- (MS-Access: Ensure referential integrity): as new records are entered in a table with a foreign key as one of the fields, it will ensure that the record will only be entered with a foreign key value that is listed in the appropriate table.



Null Values

•Refers to empty fields of a record

•Primary keys cannot be null but other fields may be null

Types Of Data Integrity In Databases

1. Table-level integrity (entity integrity):
   - Ensuring that no duplicate records exist.
   - Ensuring that no primary keys are null: MS-Access (automatic) indexed – no duplicates.
2. Relationship-level integrity (referential integrity):
   - Ensuring that relationship between a pair of tables is sound and the records in the tables are synchronized when data is entered into, updated in or deleted from either table (MS-Access: only partially implemented).
3. Field-level integrity (domain integrity):
   - Ensuring that the values in each field are valid and accurate.
   - In MS-Access this is done through input masks and validation rules.

Input Masks

- Ensures the proper format for the data entered into the database
- Example for A2: SIN number in the Employees table must be entered as:
  - <*three digits*> <space> <*three digits*> <space> <*three digits*>
- Invalid inputs:
  - Abc def ghi
  - 321 22 4234

Validation Rules

- Validation rules check the data that is entered that it is in the correct range.
- Examples for A2 (all employ the logical AND):
  - 'Employees': BirthDate
  - 'Employees': PayRate
  - 'TimeBilled': HoursWorked

Guidelines For Naming Tables

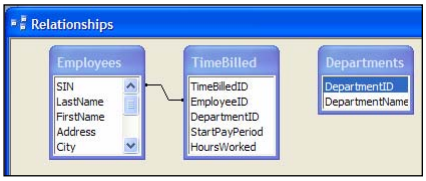
1. Create a unique and descriptive name.
2. Do not use words that convey physical characteristics or database terminology.
3. While names should be short avoid using acronyms and abbreviations unless they are well-known.
4. Do not use proper names or words that will restrict the type of data to be entered into the table.
5. Consider using the plural form of a name.
6. Avoid the use of spaces in names.

Guidelines For Naming Fields

1. Create a unique and descriptive name.
2. Create a name that accurately, clearly and unambiguously identifies the characteristic that the field represents.
3. While names should be short avoid using acronyms and abbreviations unless they are well-known.
4. Use the singular form of a name.
5. Avoid the use of spaces in names.

Relationships Between Tables

- Relationships occur when a field of one table is a foreign key in another table.



- Multiplicity: indicates how many instances of a particular item participates in the relationship:
  1. One to one
  2. One to many
  3. Many to many

Multiplicity

1. One to one relationships
   - One entity participates in the relationship from the 'left' and one entity participates in the relationship from the 'right'.
   - Person : head
   - Worker : Social Insurance Number
   - This type of relationship is rare in databases
2. One to many relationships
   - On one side of the relationship one entity participates in the relationship while on the other side: zero or more entities may participate in the relationship.
   - Person : Hair
   - Employees : TimeBilled : Departments

Multiplicity (2)

3. Many to many relationships
   - On each side of the relationship zero or more entities may participate in the relationship.
   - Students : Classes

Multiplicity (3)

3. Many to many relationships
   - This type of relationship is not directly implemented in databases:

Students table

StudentID	**StudentFirstName**	**StudentLastName**	**StudentPhone**
123456	Jamie	Smyth	553-3992
123457	Stacey	Walls	790-3992
123458	Angel	Lam	551-4993

Classes table

ClassName	***ClassNumber***	**ClassDescription**
CPSC	*203*	Introduction to Computers
CPSC	*231*	Introduction to Computer Science I
CPSC	*233*	Introduction to Computer Science II

Multiplicity (4)

3. Many to many relationships
   - Typically implemented as two one to many relationships in databases:

Students table

StudentID	StudentFirstName	…
123456	Jamie	
123457	Stacey	

Classes table

ClassName	ClassNumber	…
CPSC	203	
CPSC	231	

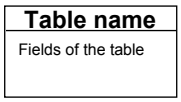
Registrations table (linking table)

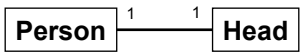
StudentID	ClassName	Class-Number
123450	ENGL	201
123457	CPSC	203
123460	MATH	271

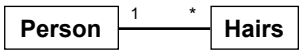
James Tam

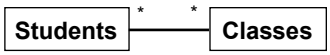
Diagrammatically Representing Relationships

Entity-Relation diagrams (E-R Diagrams or E.R.D.'s): show relationships between tables as well as any enforced rules on multiplicity:









James Tam

Retrieving Data Via Queries

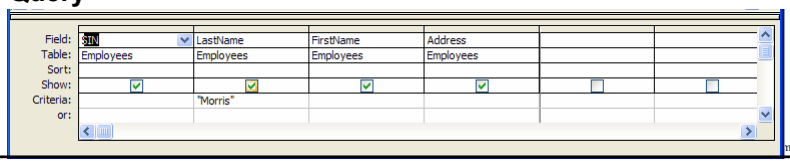
- Data retrieval occurs through the use of 'queries':
  - A query is a question asked of the data in the database.
  - Typically worded to show only the parts of the database for which the answer to the question is true.
  - Example from A2: What is the SIN, name and pay rate of every employee in the Employees Table:

Employee Pay Rates : Select Query

SIN	LastName	FirstName	PayRate
123 115 323	Simcox	Cole	30
123 456 789	Smith	John	20
371 988 812	Carswell	Mary	30
413 754 621	Kennedy	Leon	30
444 638 047	Redfield	Claire	35

  - Example: What employees have the last name of Morris?

Query



Retrieving Data Via Queries (2)

  - Example: What employees have the last name of Morris?

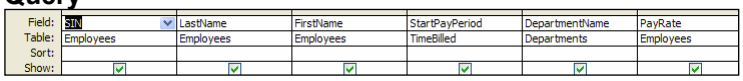
Result of query

Morris : Select Query

SIN	LastName	FirstName	Address
666 666 666	Morris	Heather	7 Luckstone Dr

- Queries can search multiple tables:
  - Example from A2: What is the gross pay of employees (3 tables searched)?

Query



Result of the query

SIN	LastName	FirstName	StartPayPeriod	DepartmentName	PayRate	HoursWorked	Gross Pay
123 115 323	Simcox	Cole	10/1/2007	Human Resources	30	40	1200
123 456 789	Smith	John	10/1/2007	Human Resources	20	40	800
371 988 812	Carswell	Mary	10/1/2007	Human Resources	30	40	1200
413 754 621	Kennedy	Leon	10/1/2007	Marketing	30	50	1500
444 638 047	Redfield	Claire	10/1/2007	Marketing	35	50	1750

Logical Operations

Operation	Description
AND	• All conditions must be true for the result to be true. • If any condition is false then the entire result is false.
OR	• All conditions must be false for the result to be false. • If any condition is true then the entire result is true.

Logical Comparisons

Operator	Description
<	Less than
<=	Less than or equal to
>	Greater than
>=	Greater than or equal to
<>	Not equal to

Logic And Queries

- Logical operators and logical comparisons can be performed during queries.
  - Examples: Which employees have the last name of 'Morris' or 'Mason'?

Query

Field:	SIN	LastName	FirstName	Address
Table:	Employees	Employees	Employees	Employees
Sort:				
Show:	☑	☑	☑	☑
Criteria:		"Morris" OR "Mason"		

Result of the query

Morris OR Mason : Select Query

SIN	LastName	FirstName	Address
666 666 667	Mason	Harry	7 Luckstone Dr
666 666 666	Morris	Heather	7 Luckstone Dr

Logic And Queries (2)

- Ranges can be specified during the query.
  - Example from A2: Which employees have a gross pay on their time card that's less than $300 or greater than $3,000 (inclusive)?

Query

Field:	SIN	LastName	FirstName	StartPayPeriod	PayRate	HoursWorked	GrossPay: [PayRate
Table:	Employees	Employees	Employees	TimeBilled	Employees	TimeBilled	
Sort:							
Show:	☑	☑	☑	☑	☑	☑	☑
Criteria:							<=300 Or >=3000
or:							

Result of the query

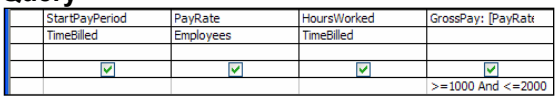
Employees with unusual pay : Select Query

SIN	LastName	FirstName	StartPayPeriod	PayRate	HoursWorked	GrossPay
456 889 123	Flanders	Ned	10/1/2007	50	80	4000
456 438 624	Lemoy	Leonard	10/1/2007	100	60	6000
620 451 097	Williams	Amanda	10/8/2007	20	10	200

Logic And Queries (3)

- Ranges can be specified during the query.
  - Example: Which employees have a gross pay within the range of $1,000 - $2000 (inclusive) on one of their timecards?

Query



	StartPayPeriod	PayRate	HoursWorked	GrossPay: [PayRate
	TimeBilled	Employees	TimeBilled	
	☑	☑	☑	☑
				>=1000 And <=2000

Result of the query

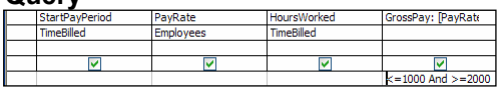
Employees with pay $1K - $2K : Select Query

SIN	LastName	FirstName	StartPayPeriod	PayRate	HoursWorked	GrossPay
456 789 123	Cartman	Eric	10/1/2007	20	80	1600
670 380 456	Edgar	Maureen	10/1/2007	50	40	2000
413 754 621	Kennedy	Leon	10/1/2007	30	50	1500
666 666 667	Mason	Harry	10/1/2007	30	50	1500
444 638 047	Redfield	Claire	10/1/2007	35	50	1750
123 115 323	Simcox	Cole	10/1/2007	30	40	1200
456 789 124	Simpson	Homer	10/1/2007	20	60	1200
666 666 668	Sunderland	James	10/1/2007	25	60	1500
371 988 812	Carswell	Mary	10/1/2007	30	40	1200

Logic And Queries (4)

- Take care not to specify queries that can never be true!
- This will result in an "Empty Query", a query that yields no results.
  - Example: Which employees have a gross pay lower than $1,000 AND higher than $2,000 on one of their time cards?

Query



	StartPayPeriod	PayRate	HoursWorked	GrossPay: [PayRate
	TimeBilled	Employees	TimeBilled	
	☑	☑	☑	☑
				<=1000 And >=2000

Result of the (empty) query

Employees with pay les than $1K AND greater than $2K : Select Query

SIN	LastName	FirstName	StartPayPeriod	PayRate	HoursWorked	GrossPay

Wav file from "The Simpson" © Fox

Using The Wildcard In Queries

- The 'wildcard' character can stand for any number of characters in the position that it's placed:
  - Example queries that follow will be in the Employees table:

Employees : Table

SIN	LastName	FirstName	Address	City	Province	PostalCode
123 115 323	Simcox	Cole	311 Ocean View Drive	Vancouver	British Columbia	T1N-4N9
123 456 789	Smith	John	123 Peanut Lane	Calgary	Alberta	T1N-3N4
371 988 812	Carswell	Mary	425 Remington Ave	Calgary	Alberta	T3N-7N4
413 754 621	Kennedy	Leon	808, 4900 Wildman A	Racoon City	Alberta	T2S-1M0
444 638 047	Redfield	Claire	653 Wildpark Place	Racoon City	Alberta	T2S-1M0
456 438 624	Lemoy	Leonard	55 Logic Way	Vulcan	Alberta	VS1-3N3
456 789 123	Cartman	Eric	456 Lynchview Road	Southpark	Alberta	S0S-9A9
456 789 124	Simpson	Homer	59 Evergreen Terrace	Springfield	Alberta	N1E-7X6
456 889 123	Flanders	Ned	60 Evergreen Terrace	Springfield	Alberta	N1E-7X6
620 451 097	Williams	Amanda	25 Rodeo Drive	Edmonton	Alberta	V6N-6N5
638 666 670	Cartland	Douglas	1109, 4944 Dalworth	Silent Hill	Alberta	S6N-9X9
666 666 666	Morris	Heather	7 Luckstone Dr	Silent Hill	Alberta	T3A-3H1
666 666 667	Mason	Harry	7 Luckstone Dr	Silent Hill	Alberta	T3A-3H1
666 666 668	Sunderland	James	7 Heartbroken Ave	Silent Hill	Alberta	T3A-2E6
666 666 669	Wolf	Claudia	66 Twisted View	Silent Hill	Alberta	T1N-3O4
670 380 456	Edgar	Maureen	300, Lockinvar Road	Calgary	Alberta	T4P-3N9

Using The Wildcard In Queries

- Examples:
  - Which employees have a last name that begins with 'm'?

LastName	FirstName
Mason	Harry
Morris	Heather

  - Which employees have a last name ends with 's'?

LastName	FirstName
Flanders	Ned
Morris	Heather
Williams	Amanda

  - Which employees have the letter 'a' anywhere in their first name?

LastName	FirstName
Cartland	Douglas
Edgar	Maureen
Lemoy	Leonard
Mason	Harry
Morris	Heather
Redfield	Claire
Sunderland	James
Williams	Amanda
Wolf	Claudia
Carswell	Mary

Single Character Wildcard

- The '?' stands for a single character wildcard:
  - Querying the following table

EmployeesVersion2 : Table

LastName	FirstName	SIN
Williams	Robert	123 456 789
Scalisce	Rita	111 222 444
Lam	Angel	222 222 222
Nelson	Roberta	333 333 333
Ashland	Renert	456 789 999

  - Which employees have the following string of characters in their first name: *<R> <any character> <B> <any number of characters>*

R?B : Select Query

LastName	FirstName
Williams	Robert
Nelson	Roberta

Object-Oriented Databases

- Developed because relational databases sometimes cannot handle the storage of complex data (e.g., images, videos)

Relational database (tables)

FirstName	LastName	Address	…
Jessica	Gravowski	123 Summerset Road	
Stacey	Walls	#80 Sunvalley Way	
:	:	:	

Table

Object-Oriented database (objects)

FirstName	LastName	Address	…
Jessica	Gravowski	123 Summerset Road	
Stacey	Walls	#80 Sunvalley Way	
:	:	:	

Data

CheckCredit (action)	ProcessOrder (action)	UpdateAddress (action)

Actions

Object

You Should Now Know

- How a database is broken down into tables and how tables are broken down into it's component parts
- What are the type of tables and the purpose of each
- What is the purpose of a primary key
- What is a foreign key
- What is the purpose of creating a table with foreign keys
- What is a null value
- What are forms of data integrity in databases
- How is the integrity of data in database provided through input masks and validation rules

You Should Now Know (2)

- Guidelines for naming tables and the fields of the tables
- What are the three relationships that may exist between tables and how they differ
- How is a many-to-many relationship typically implemented in a database
- The ERD representation of databases
- How to form different queries in order to retrieve data from a database
- What is an empty query
- How wildcards can be used in queries
- What is an Object-Oriented database and how it differs from a relational database