Technology in Action

Chapter 11

Behind the Scenes:
Databases and Information Systems
Chapter Topics

- Databases and their uses
- Database components
- Types of databases
- Database management systems
- Relational databases
- Data warehouses and data marts
- Information systems
- Data mining
Databases

- Electronic collections of related data
- Help us organize data
- Turn data into information
Advantages of Using Databases

- Store and retrieve large quantities of information
- Enable information sharing
- Provide data centralization
- Promote data integrity
- Allow for flexible use of data
Database Terminology

- **Field**
  - Category of information displayed in columns
  - Each field has a field name
Database Terminology

• Data types
  – Text
  – Numeric
  – Computational
  – Date
  – Memo
  – Object
  – Hyperlink
### Database Terminology

- **Record**
  - A group of related fields

<table>
<thead>
<tr>
<th>Customer Num</th>
<th>Name</th>
<th>Address</th>
<th>City</th>
<th>State</th>
<th>Zip Code</th>
<th>Balance</th>
<th>Technic</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT23</td>
<td>Atlas Repair</td>
<td>220 Beard</td>
<td>Kady</td>
<td>TN</td>
<td>42514</td>
<td>$335.00</td>
<td>203</td>
</tr>
<tr>
<td>AZ01</td>
<td>AZ Auto</td>
<td>412 Beechwood</td>
<td>Conrad</td>
<td>TN</td>
<td>42547</td>
<td>$300.00</td>
<td>210</td>
</tr>
<tr>
<td>BL35</td>
<td>Blarton Shoes</td>
<td>443 Cheddar</td>
<td>Kady</td>
<td>TN</td>
<td>42514</td>
<td>$290.00</td>
<td>210</td>
</tr>
<tr>
<td>CJ45</td>
<td>C Joe Diner</td>
<td>37 Fletcher</td>
<td>Carlton</td>
<td>TN</td>
<td>52764</td>
<td>$0.00</td>
<td>214</td>
</tr>
<tr>
<td>CM80</td>
<td>Cramden Co.</td>
<td>234 Fairawn</td>
<td>Conrad</td>
<td>TN</td>
<td>42546</td>
<td>$355.00</td>
<td>203</td>
</tr>
<tr>
<td>HI25</td>
<td>Hill Crafts</td>
<td>245 Beard</td>
<td>Kady</td>
<td>TN</td>
<td>42514</td>
<td>$334.00</td>
<td>210</td>
</tr>
<tr>
<td>KL50</td>
<td>Klean n Dri</td>
<td>379 Stout</td>
<td>Carlton</td>
<td>TN</td>
<td>52764</td>
<td>$365.00</td>
<td>214</td>
</tr>
<tr>
<td>MC10</td>
<td>Moss Carpet</td>
<td>109 Fletcher</td>
<td>Carlton</td>
<td>TN</td>
<td>52764</td>
<td>$398.00</td>
<td>203</td>
</tr>
<tr>
<td>PV83</td>
<td>Prime Vdeo</td>
<td>734 Lanton</td>
<td>Conrad</td>
<td>TN</td>
<td>42547</td>
<td>$0.00</td>
<td>214</td>
</tr>
<tr>
<td>SE05</td>
<td>Serve Mfg Co.</td>
<td>979 Redfern</td>
<td>Kady</td>
<td>TN</td>
<td>42515</td>
<td>$343.00</td>
<td>210</td>
</tr>
<tr>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$0.00</td>
<td></td>
</tr>
</tbody>
</table>
Database Terminology

• Table
  – A group of related records
Database Terminology

- **Primary key**
  - A field value unique to a record

<table>
<thead>
<tr>
<th>SS#</th>
<th>Last Name</th>
<th>First Name</th>
<th>Home Address</th>
<th>City</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>234567891</td>
<td>Chan</td>
<td>Li</td>
<td>123 Main Street</td>
<td>Tuba City</td>
<td>NV</td>
</tr>
<tr>
<td>456789123</td>
<td>Coyle</td>
<td>Diane</td>
<td>745 Station Drive</td>
<td>Springfield</td>
<td>MA</td>
</tr>
<tr>
<td>987654321</td>
<td>Evans</td>
<td>Jennifer</td>
<td>312 Brookside Road</td>
<td>Springfield</td>
<td>MA</td>
</tr>
<tr>
<td>123456789</td>
<td>Evans</td>
<td>Jennifer</td>
<td>123 Oak Street</td>
<td>Gotham City</td>
<td>PA</td>
</tr>
<tr>
<td>567891234</td>
<td>Lopez</td>
<td>Donald</td>
<td>3421 Lincoln Court</td>
<td>Spalding</td>
<td>ND</td>
</tr>
<tr>
<td>345678912</td>
<td>Wallace</td>
<td>William</td>
<td>654 Front Street</td>
<td>Locust Glen</td>
<td>MI</td>
</tr>
</tbody>
</table>
Database Types

• Relational databases
  – Organize data in a table
  – Link tables to each other through their primary keys

• Object-oriented databases
  – Stores data in objects
  – Handles unstructured data

• Multidimensional databases
  – Stores data in multiple dimensions
  – Can easily be customized
  – Process data much faster
Database Management Systems (DBMS)

- Application software designed to capture and analyze data
- Four main operations of a DBMS are:
  1. Creating databases and entering data
  2. Viewing and sorting data
  3. Extracting data
  4. Outputting data
1. Creating Databases and Entering Data

- Create field names:
  - Identify each type of data
  - Data dictionary
Creating Databases and Entering Data (cont)

- Create individual records:
  - Key-in
  - Import

<table>
<thead>
<tr>
<th>Mailing List ID</th>
<th>First Name</th>
<th>Last Name</th>
<th>Address</th>
<th>City</th>
<th>State</th>
<th>Postal Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>Jason</td>
<td>Alexander</td>
<td>123 Main Street</td>
<td>Anywhere</td>
<td>USA</td>
<td>12345</td>
</tr>
<tr>
<td># (AutoNumber)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes on the form provide additional guidance for users.
The blank boxes visually indicate the length of the data fields.
Data Validation

• Validation
  – Process of ensuring data entered into the database is correct (or at least reasonable) and complete

• Validation rules
  – Range check
  – Completeness check
  – Consistency check
  – Alphabetic/numeric checks
2. Viewing and Sorting Data

- Browse through records
- Sort records by field name

Before sort

<table>
<thead>
<tr>
<th>Social Security</th>
<th>First Name</th>
<th>Last Name</th>
<th>Address</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>123-45-6789</td>
<td>Jennifer</td>
<td>Evans</td>
<td>123 Oak Street</td>
<td>Gotham City</td>
</tr>
<tr>
<td>234-56-7891</td>
<td>Li</td>
<td>Chan</td>
<td>123 Main Street</td>
<td>Tuba City</td>
</tr>
<tr>
<td>345-67-8912</td>
<td>William</td>
<td>Wallace</td>
<td>664 Front Street</td>
<td>Locust Glen</td>
</tr>
<tr>
<td>456-78-9123</td>
<td>Diane</td>
<td>Coyle</td>
<td>745 Station Drive</td>
<td>Springfield</td>
</tr>
<tr>
<td>543-21-9876</td>
<td>Austin</td>
<td>Knowland</td>
<td>345 Reston Court</td>
<td>Red Ridge</td>
</tr>
<tr>
<td>654-32-1987</td>
<td>Donald</td>
<td>Lopez</td>
<td>3421 Lincoln Court</td>
<td>Spalding</td>
</tr>
<tr>
<td>765-43-2198</td>
<td>Susan</td>
<td>Hudson</td>
<td>237 Explorer Circle</td>
<td>Booth Bay</td>
</tr>
<tr>
<td>876-54-3219</td>
<td>Michelle</td>
<td>Swanson</td>
<td>1234 North Street</td>
<td>Red Ridge</td>
</tr>
<tr>
<td>387-55-4321</td>
<td>Kevin</td>
<td>Po</td>
<td>1632 Delsey Drive</td>
<td>Blue Ridge</td>
</tr>
</tbody>
</table>

Sorting Options: Ascending (selected) or Descending.

After sort

<table>
<thead>
<tr>
<th>Social Security</th>
<th>First Name</th>
<th>Last Name</th>
<th>Address</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>234-56-7891</td>
<td>Li</td>
<td>Chan</td>
<td>123 Main Street</td>
<td>Tuba City</td>
</tr>
<tr>
<td>456-78-9123</td>
<td>Diane</td>
<td>Coyle</td>
<td>745 Station Drive</td>
<td>Springfield</td>
</tr>
<tr>
<td>123-45-6789</td>
<td>Jennifer</td>
<td>Evans</td>
<td>123 Oak Street</td>
<td>Gotham City</td>
</tr>
<tr>
<td>654-32-1987</td>
<td>Susan</td>
<td>Hudson</td>
<td>237 Explorer Circle</td>
<td>Booth Bay</td>
</tr>
<tr>
<td>543-21-9876</td>
<td>Austin</td>
<td>Knowland</td>
<td>345 Reston Court</td>
<td>Red Ridge</td>
</tr>
<tr>
<td>876-54-3219</td>
<td>Fortune</td>
<td>Liewiski</td>
<td>141 Central Blvd</td>
<td>Wyandott</td>
</tr>
<tr>
<td>387-554-3219</td>
<td>Donald</td>
<td>Lopez</td>
<td>3421 Lincoln Court</td>
<td>Spalding</td>
</tr>
<tr>
<td>987-65-4321</td>
<td>Kevin</td>
<td>Po</td>
<td>1632 Delsey Drive</td>
<td>Blue Ridge</td>
</tr>
<tr>
<td>765-43-2198</td>
<td>Michelle</td>
<td>Swanson</td>
<td>1234 North Street</td>
<td>Red Ridge</td>
</tr>
<tr>
<td>345-67-8912</td>
<td>William</td>
<td>Wallace</td>
<td>654 Front Street</td>
<td>Locust Glen</td>
</tr>
</tbody>
</table>
3. Extracting or Querying Data

- **Query**
  - A question or inquiry
  - Provides records based on criteria
  - Structured query language (SQL)

```
SELECT [Student Info].FirstName, [Student Info].LastName, [Student Info].Address, [Student Info].City, [Student Info].State, [Student Info].Zip Code
FROM [Student Info];
```
4. Outputting Data

- Reports:
  - Printed
  - Summary data reports
- Export data
Relational Database Operations

- Relational databases organize data into tables based on logical groupings.
- Relationships are links between tables with related data.
- Common fields between tables need to exist.
- Normalization of data (recording data once) reduces data redundancy.
Data Storage

• Data warehouses
  – A large scale repository of data
  – Organizes all the data related to an organization
  – Data is organized by subject
Populating Data Warehouses

- **Source data**
  - **Internal sources**
    - Company databases, etc.
  - **External sources**
    - Suppliers, vendors, etc.
  - **Customers or Web site visitors**
    - Clickstream data
Data Staging

• Data staging
  – Extract data from source
  – Reformat the data
  – Store the data

• Software programs/procedures created to extract the data and to reformat it for storage
Data Marts

• Small slices of data
• Data of a single department
Data Warehouse

Source Data (Internal & External Sources)
- Supplier Database
- Customer Order Database
- Billing Database

Web Users
- Clickstream Data

OLAP Query

Information Systems (To Extract Data)

Data Staging
Data Reformatting

Data Warehouse DMBS
Reformatted Data Stored

Data Marts
- Sales Department Data Mart
- Customer Service Data Mart
Managing Data: Information Systems

• Information systems
  – Software-based solutions used to gather and analyze information

• Functions performed by information systems include
  – Acquiring data
  – Processing data into information
  – Storing data
  – Providing output options
Information Systems Categories

- Office support
- Transaction processing
- Management
- Decision support
Office Support Systems (OSS)

- Assist employees in day-to-day tasks
- Improve communications
- Microsoft Office
Transaction Processing Systems (TPS)

- Keeps track of everyday business activities
- Batch processing
- Real-time processing
Management Information Systems (MIS)

• Provides timely and accurate information for managers to make business decisions

• Detail report:
  – Transactions that occur during a period of time

• Summary report:
  – Consolidate detailed data

• Exception report:
  – Show unusual conditions
Decision Support Systems (DSS)

• Help managers develop solutions for specific problems
Data Mining

- Process by which data are analyzed and investigated
- Objective is to spot patterns or trends within the data
Data Mining Methods

- **Classification**
  - Define data classes
- **Estimation**
  - Assign a value to data
- **Affinity grouping or association rules**
  - Determine which data goes together
- **Clustering**
  - Organize data into subgroups
- **Description and visualization**
  - Provides a clear picture of what is happening