Test Your Tech

A spread sheet:
A. Only happens on laundry day.
B. Is covered with food during holiday meals.
C. Helps answer "what-if" questions.

11/30/2009  D.A. Clements, UW Information School

Announcement

• Software for rest of quarter
  • Microsoft Excel in Lab 10
  • Microsoft Access for the rest of the labs and Project 3
    • Microsoft Access—PC’s only
    • Mac users will have to use the labs on campus
    • No Mac equivalent

11/30/2009  D.A. Clements, UW Information School

Announcements

• Free copy of Access, Windows 7, etc., for educational/academic use:
  • Links on Computing page on Course Web site
  • Search for CSE or INFO to find the link on the page
  • Username is your full UW email address
  • Password is different!
    • Click on "send a reminder"
    • Check wherever your email forwards to

11/30/2009  D.A. Clements, UW Information School

Announcements

• Readings
  • Today—Ch 15
  • Wednesday—Ch 16

11/30/2009  D.A. Clements, UW Information School

Announcements

• Project 2B due Wednesday night at 10pm
• Drop-In Labs
  • Tuesday 8:30am MGH 430
  • Tuesday 5pm MGH 430
  • Wednesday CLUE Tutoring 7pm MGH 058
    • 2 pts extra credit per CLUE session

11/30/2009  D.A. Clements, UW Information School
Announcements

- Lab 10 due Friday night at 10pm
- Drop-In Labs
  - Thursday 8:30am MGH 430
  - Friday 1:30pm MGH 430

Drop-In Labs

- Thursday 8:30am MGH 430
- Friday 1:30pm MGH 430

Tight deadlines for rest of course
- No extensions
- No lab time scheduled for Project 3A

Where we've been...

- Unit I—Connections
  - Hardware, networks, protocols, Internet, Web, building Web pages
- Unit II—Programming
  - Concepts common to all programming languages
- Unit III—Data
  - Storage, retrieval, transfer

Unit III: Data Storage, Transfer, and Retrieval

Keeping your private information private and secure

D.A. Clements

- Storage
  - Format—physical and logical
- Retrieval
  - The information you need when you need it
- Transfer
  - Between people, departments, organizations
  - Media—spreadsheets, databases, XML

Spreadsheets

- Spreadsheets are a powerful abstraction for organizing data and computation

D.A. Clements, UW Information School

11/30/2009
A spreadsheet is a 2-dimensional array of cells...it’s 3D with multiple worksheets
- Rows or columns represent a single data type
  - They will be operated on similarly, so that’s easy to do
  - Adding more data of the same type means adding more rows or columns
  - Often spreadsheets contain numbers, but text-only spreadsheets are useful, too

Spreadsheets are not so unusual
- The position (row/column) names the data, as with memory locations, variables, forms, etc.
- Operating on all elements of a column (or row) is an iteration, though not the World Famous Iteration
- Setting a cell to a formula is an (unevaluated) assignment statement with cells as variables
- The formula is an expression
- Functions are built-in expressions

Familiar Terminology
- column heading
- row name
- cell
- formula
- column name
- referenced cell E2

Familiar Terminology
- column heading
- row name
- cell
- formula
- column name
- referenced cell E2
Formulas

The data in a spreadsheet can be manipulated using formulas.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>Common Name</strong></td>
<td><strong>Distance (km)</strong></td>
<td><strong>Body Length (m)</strong></td>
</tr>
<tr>
<td>2</td>
<td>Swainson's Hawk</td>
<td>13500</td>
<td>0.621</td>
</tr>
<tr>
<td>3</td>
<td>Wheatear</td>
<td>15500</td>
<td>0.16</td>
</tr>
<tr>
<td>4</td>
<td>Willow Warbler</td>
<td>15500</td>
<td>0.11</td>
</tr>
<tr>
<td>5</td>
<td>Short-tailed Shy</td>
<td>12500</td>
<td>0.41</td>
</tr>
<tr>
<td>6</td>
<td>Long-Tailed Shy</td>
<td>16000</td>
<td>0.31</td>
</tr>
<tr>
<td>7</td>
<td>Arctic Tern</td>
<td>19200</td>
<td>0.35</td>
</tr>
</tbody>
</table>

The value in D2 (selected cell) is the value in B2 times 0.621...the result is shown but the cell has the formula.

Filling Replicates Formulas

Fill is a spreadsheet shortcut for copy-and-paste.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>Common Name</strong></td>
<td><strong>Distance (km)</strong></td>
<td><strong>Body Length (m)</strong></td>
</tr>
<tr>
<td>2</td>
<td>Swainson's Hawk</td>
<td>13500</td>
<td>0.621</td>
</tr>
<tr>
<td>3</td>
<td>Wheatear</td>
<td>15500</td>
<td>0.16</td>
</tr>
<tr>
<td>4</td>
<td>Willow Warbler</td>
<td>15500</td>
<td>0.11</td>
</tr>
<tr>
<td>5</td>
<td>Short-tailed Shy</td>
<td>12500</td>
<td>0.41</td>
</tr>
<tr>
<td>6</td>
<td>Long-Tailed Shy</td>
<td>16000</td>
<td>0.31</td>
</tr>
<tr>
<td>7</td>
<td>Arctic Tern</td>
<td>19200</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Grab the fill tab and pull in the direction to be pasted.

A Powerful Translation

- The graphic shows the equations in the cells with the translation:
  - The row changes going down but the column doesn't.

Relative & Absolute Addressing

- References to cells happens in two ways: Relative and Absolute (with $)
  - $F2 relative column, absolute row
  - $F$2 absolute column, absolute row
  - $F$2 absolute column, absolute row

Relative references change when pasted/filled; absolute references do not!

An Example

Creating a discount table uses both relative and absolute refs
- Consider store credit of $1 per $10 spent
- $3 store credit for every 2 CDs (1 earns $1)

<table>
<thead>
<tr>
<th>(CDs Purchased)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spent</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10</td>
<td>$2.00</td>
<td>$4.00</td>
<td>$6.00</td>
<td>$8.00</td>
<td>$10.00</td>
<td>$12.00</td>
<td>$14.00</td>
<td>$16.00</td>
</tr>
<tr>
<td>$20</td>
<td>$4.00</td>
<td>$8.00</td>
<td>$12.00</td>
<td>$16.00</td>
<td>$20.00</td>
<td>$24.00</td>
<td>$28.00</td>
<td>$32.00</td>
</tr>
<tr>
<td>$30</td>
<td>$6.00</td>
<td>$12.00</td>
<td>$18.00</td>
<td>$24.00</td>
<td>$30.00</td>
<td>$36.00</td>
<td>$42.00</td>
<td>$48.00</td>
</tr>
<tr>
<td>$40</td>
<td>$8.00</td>
<td>$16.00</td>
<td>$24.00</td>
<td>$32.00</td>
<td>$40.00</td>
<td>$48.00</td>
<td>$56.00</td>
<td>$64.00</td>
</tr>
<tr>
<td>$50</td>
<td>$10.00</td>
<td>$20.00</td>
<td>$30.00</td>
<td>$40.00</td>
<td>$50.00</td>
<td>$60.00</td>
<td>$70.00</td>
<td>$80.00</td>
</tr>
<tr>
<td><strong>Earned</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10</td>
<td>$1.00</td>
<td>$2.00</td>
<td>$3.00</td>
<td>$4.00</td>
<td>$5.00</td>
<td>$6.00</td>
<td>$7.00</td>
<td>$8.00</td>
</tr>
<tr>
<td>$20</td>
<td>$2.00</td>
<td>$4.00</td>
<td>$6.00</td>
<td>$8.00</td>
<td>$10.00</td>
<td>$12.00</td>
<td>$14.00</td>
<td>$16.00</td>
</tr>
<tr>
<td>$30</td>
<td>$3.00</td>
<td>$6.00</td>
<td>$9.00</td>
<td>$12.00</td>
<td>$15.00</td>
<td>$18.00</td>
<td>$21.00</td>
<td>$24.00</td>
</tr>
<tr>
<td>$40</td>
<td>$4.00</td>
<td>$8.00</td>
<td>$12.00</td>
<td>$16.00</td>
<td>$20.00</td>
<td>$24.00</td>
<td>$28.00</td>
<td>$32.00</td>
</tr>
<tr>
<td>$50</td>
<td>$5.00</td>
<td>$10.00</td>
<td>$15.00</td>
<td>$20.00</td>
<td>$25.00</td>
<td>$30.00</td>
<td>$35.00</td>
<td>$40.00</td>
</tr>
</tbody>
</table>

A cell is based on first column, top row data in that row and column...must mix relative and absolute references.
Series

- Another handy property of fill is that it can make a series based on constants
  - Fill Sunday => Monday, Tuesday, Wed...
  - Fill 22 Feb => 23 Feb, 24 Feb, 25 Feb...
- More generally
  - Series fill will even count using a constant
  - Counting by odd sizes: give 1st two items

Tables

- Familiar format of rows and columns
- Can work directly with the data
- "What-If" scenarios
- Involved computations like taxes
- Storing lists

Microsoft reports that...

- 70% of Excel users use it like a database

Advantages of Spreadsheets

- Familiar format of rows and columns
- Can work directly with the data
- "What-If" scenarios
- Involved computations like taxes
- Storing lists

Databases are better...

- When you...
  - Have a lot of spreadsheets
  - Need to pass data back and forth between spreadsheets
  - Scroll a lot to find answers
  - Have a lot of repetitious data, like
    - Many contacts at same company with
    - Repeated company address for each one
Databases are better

- No repetition
- Can search for exactly the data you need
- Solves the problem of information overload

Database Advantage

- You can
  - Save a query for later
  - Use over and over and over again
  - Edit the query
  - Copy the query
  - Format an attractive report that prints every time you run the query

Example

- List all students who received a "B"
  - Spreadsheet
    1. Sort the data
    2. Scroll (and scroll) to find those in the "B" range
  - Database
    1. Query for students who received a "B"
      - Results: All the students and only the students who received a "B"

The Database Disadvantage

- It takes time to set up the database and make sure it's working properly
- Spreadsheets are easier, faster to set up
  - If they're small

How do you decide?

- Do changes made in one spreadsheet force you to make changes in another?
- Do you have several spreadsheets containing similar information (such as separate sheets with inventory for Dallas, D.C., and Detroit)?
- Do you want some data to be hidden from some users?
Move to a database if…

- You answered "yes" to at least 2 of these questions

Next lecture…

- We'll continue to look at data storage, transfer, and retrieval
- Read Chapter 16