



USAID
FROM THE AMERICAN PEOPLE

Regulatory and Energy Assistance

Navigant Consulting, Inc.

Training in Designing and Building Models Using Microsoft Excel

For the FERK Tariff Staff

May 2010

Bosnia and Herzegovina (BiH) Regulatory and Energy Assistance Project (REAP)

USAID Contract No. EPP-I-00-03-00004-00, Task Order 5

Implemented by Advanced Engineering Associates International, Inc. as USAID Contractor

This presentation is made possible by support from the American People sponsored by the United States Agency for International Development (USAID). The contents are the sole responsibility of the author/s and do not necessarily reflect the views of USAID or the United States Government.

Table of Contents

1	Introduction
2	The Model Lifecycle
3	Model Scoping and Specification
4	Model Design
5	Model Construction
6	Model Documentation
7	Model Testing



Table of Contents

1	Introduction
2	The Model Lifecycle
3	Model Scoping and Specification
4	Model Design
5	Model Construction
6	Model Documentation
7	Model Testing



Goals of the Training

- Provide the FERK Tariff Staff with an enhanced knowledge of Microsoft Excel
- Provide a systematic approach to the modeling process which would ensure a high level of accuracy, transparency and easily change the model as needed.

Approach

- Multiple sessions.
- Assignments which will be based upon the partially completed tariff model.
- Work in groups.



Protocol

- What happens if we find a mistake or problem?



Blame Ivan!



USAID
FROM THE AMERICAN PEOPLE

Regulatory and Energy Assistance

Table of Contents

1	Introduction
2	The Model Lifecycle
3	Model Scoping and Specification
4	Model Design
5	Model Construction
6	Model Documentation
7	Model Testing



Objectives

By participation in this Module, students should be able to:

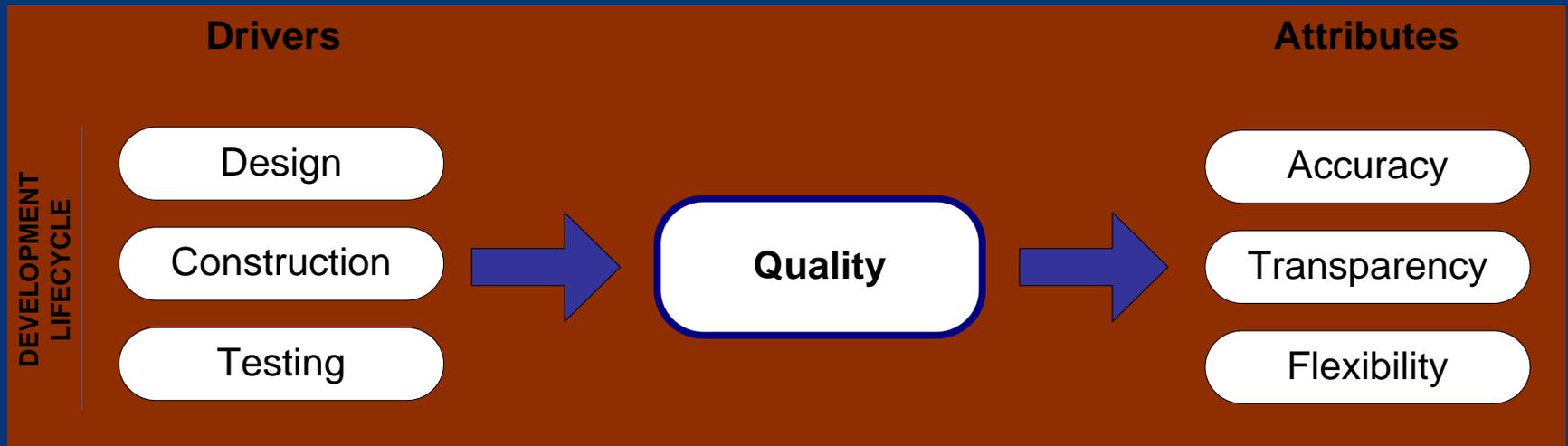
- Understand the importance of a structured, temporal approach for spreadsheets

Spreadsheet Challenges

- Analysts often have incentives to produce results in the shortest amount of time
- A single consultant is responsible for the design, construction, and testing
- Scoping, design, documentation and testing, which should comprise up to 75% of the development cycle, are given less attention
- The result is often a hastily built model with poor structure and avoidable errors

Spreadsheet Quality

- Spreadsheet Quality is a function of efforts expended in the development process, designing, constructing, documenting and testing the model
- Quality is primarily reflected in three model attributes: accuracy, transparency, and flexibility



Spreadsheet Quality *(Continued)*

- Poor development practices can lead to:
 - Obvious and latent errors
 - Opaqueness
 - Rigidity

The Model Lifecycle



- Scoping
- Specification
- Blueprint

- Construction/
documentation of
model per design

- Auditing tools
- Execution testing
- Cell inspection
- Testing programs

Questions



USAID
FROM THE AMERICAN PEOPLE

Regulatory and Energy Assistance

Assignment

- Take 10 minutes and with your team outline what is required to:
 - Scope
 - Outline
 - Document
 - Test
- The FERK Tariff Model

Table of Contents

1	Introduction
2	The Model Lifecycle
3	Model Scoping and Specification
4	Model Design
5	Model Construction
6	Model Documentation
7	Model Testing



Objectives

By participation in this Module, students should be able to:

- Explain the importance of Model Design in the Model Lifecycle
- Use tools such as Bubble Charts, Variable Outlines, Calculation Tables and Prototype Models
- Explain the importance of properly structuring a workbook, including arranging worksheets within that workbook

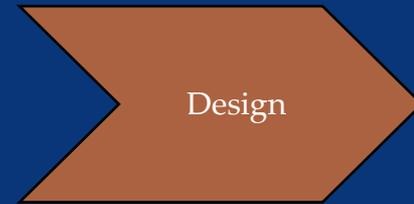
Objectives *(Continued)*

By participation in this Module, students should be able to:

- Explain data and labeling limitations
- Build models with scenario and sensitivity analysis capabilities
- Be consistent within a model as it relates to formatting, range names, formulas, organization, worksheet labeling, and documentation

Scope

Purpose: To develop a clear understanding of the problem and to propose a means by which the spreadsheet will address the problem.



Useful Techniques:

- Workshops with model sponsors and consulting team
- Written document defining model specifications
- Written documentation discussing any changes in scope

Scoping

Specification

Blueprint

Model Scope Exercise

1

What You Need To Do

- Work in teams.
- Your supervisor has sent you an e-mail, asking you to update a spreadsheet. Based only on the text of the message, and without opening the attachment, create a list of five questions you might ask your supervisor in order to confirm his/her expectations.
- You will have 5 minutes to compile your questions.

Specification

- *Purpose:* Minimize model redesigns resulting from poor initial conception.
- *Considerations:* Effective solutions will vary depending on requirements, including:
 - “In-house” vs. “Public” model
 - Intended audience



Scoping

Specification

Blueprint

Specification (Continued)

- *Useful Tools:*
 - Bubble Charts and Variable Outlines
 - Graphically link the outputs of the model to the data and calculations that produce them.
 - Calculation Tables
 - Connect model outputs and inputs through detailed formulas.
 - Prototype Models
 - Simplistic representation of the ultimate model used in development



Scoping

Specification

Blueprint

Specification Exercise

What You Need To Do

- Work in your table teams
- The instructor will assign a model specification tool to each team (Bubble Chart, Calculation Table, Prototype Model)
- Open the spreadsheet attached to the e-mail from Exercise 1, and use your assigned model specification tool to illustrate how average supply cost is calculated.
- You will have **15 minutes** to construct your charts on flipchart paper

Blueprint

- *Purpose:* Produce a blueprint for model inputs, outputs, and calculations.
- *Benefits:* Intelligent, logical structuring produces many efficiency gains:
 - Increases modeler productivity
 - Reduces model run-time
 - Increases model flexibility, avoiding costly workarounds
 - Reduces required guidance for peers and clients



Scoping

Specification

Blueprint

Assignment for Tomorrow

- Prepare a draft document which will:
 - Provide the Scope of the Tariff Design Model
 - Outline the Design of the Model using figures, charts and narratives as required
 - A “Blueprint” of the model.



Table of Contents

1	Introduction
2	The Model Lifecycle
3	Model Scoping and Specification
4	Model Design
5	Model Construction
6	Model Documentation
7	Model Testing

Inputs, Calculations, Outputs, And Switches

- ① *Inputs*: Raw data that impacts the ultimate calculation
- ② *Calculations*: Formulas that produce desired outputs
- ③ *Outputs*: Reports that reformat calculation results (can be exported)
- ④ *Switches*: Data that manipulate the inputs used in the calculation

Example, simple model (single sheet)

	④		
① Inputs	Case in Use		
	1	1	2
Escalation Factor	2.20%	2.20%	3.00%
Revenue 2007	10,000	10,000	15,000
② Calculations			
	2007	2008	2009
Revenue	10,000	10,220	10,445
③ Outputs			
	Year	Revenue (\$000)	
	2007	10,000	
	2008	10,220	
	2009	10,445	
	2010	10,758	

Inputs, Calculations, Outputs, And Switches - Layout

Typical simple model layout:

- Sheet 1:
 - Switches
 - Inputs
 - Calculations/outputs.
- Sheet 2:
 - Outputs

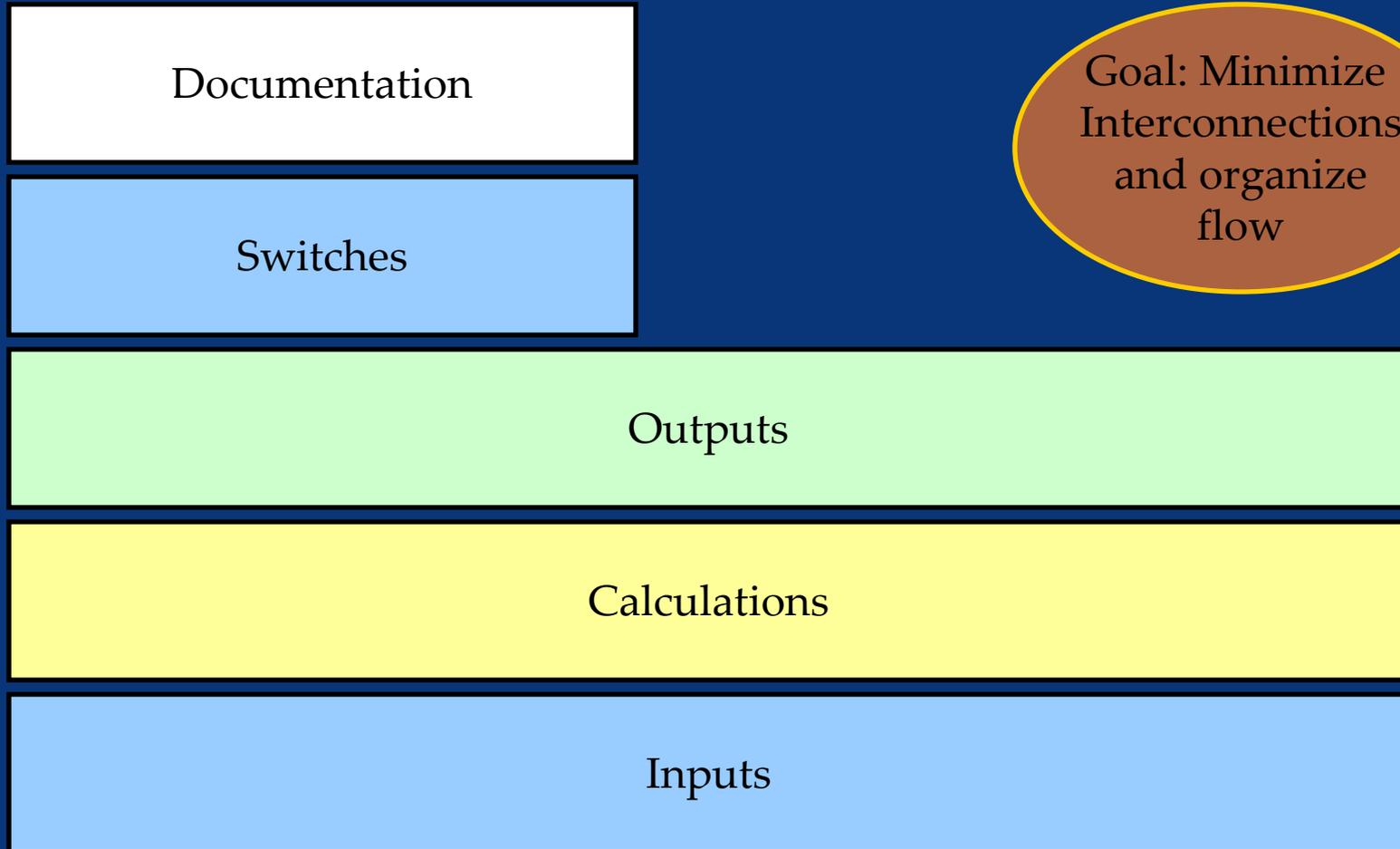
Typical complex (multi-sheet) model layout, from left to right:

- Switches
- Outputs
- Calculations
- Inputs

Documentation sheets are useful:

- describes the model's purpose
- provides the user with any other necessary instruction
- Typically located as the first sheet (left-most) in a spreadsheet

Workbook Structure – Example Layout



Exercise – Workbook Structure

3

What You Need To Do

- Work in your teams.
- In the spreadsheet, identify as many inputs as you can. List them (e.g., Sheet1!B2:E7).
- Do the same for calculations and outputs.



Scalars, Vectors, And Matrices

- Scalars: Single cell inputs (e.g., start date, interest rate)

Interest Rate	5.5%
Discount Rate	14%
Start Year	2008

Vectors: A series of related inputs with a single dimension (e.g., time)

	Jan-06	Feb-06	Mar-06	Apr-06	May-06	Jun-06
NG price (\$/Mmbtu)	9.4	9.9	8.8	9.1	10.2	11.9

- Matrices: A series of related inputs with multiple dimensions (e.g., time, plant)

Coal Prices	Jan-06	Feb-06	Mar-06	Apr-06	May-06	Jun-06
Plant 1	4.5	4.2	5.1	4.9	4.5	4.3
Plant 2	5.1	4.9	5.3	5.2	5.4	5.1
Plant 3	3.9	4.1	4.1	3.8	4.8	4.7

TIPS: Avoid adding unnecessary rows/columns within datasets
All data should be entered only once in a workbook

How Do You Reference Inputs?

- You can “Point and Click” your mouse, but this approach has shortcomings:
 - What is your input is used several times in model – “Point and Click” can lead to a incorrect reference.
 - What is your are referencing multiple elements of a table (e.g., KWH Sales by Tariff).
- “Point and Click” can lead to errors

Better Alternatives to Reference Inputs: Range Names

- Use Range Names – Especially useful for a Scalar.
- Example using the payment function



Example of Range Names

Variable	Value
Number of Payments	10
Amount	\$ 1,000
Interest	6.00%

=PMT(interest,ar

Function Arguments

PMT

Rate interest = 0.06

Nper amount = 1000

Pv number| = 10

Fv = number

Type = number

= -0.6

Calculates the payment for a loan based on constant payments and a constant interest rate.

Pv is the present value: the total amount that a series of future payments is worth now.

Formula result = -0.6

[Help on this function](#)

OK Cancel

Vertical and Horizontal Lookup Function

- References a range of values.
- Vlookup – references a range of values in a vertical vector
- Hlookup - references a range of values in a horizontal vector



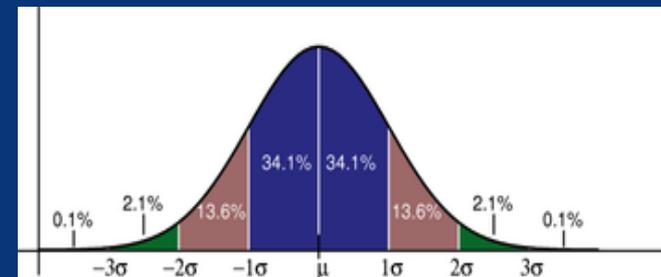
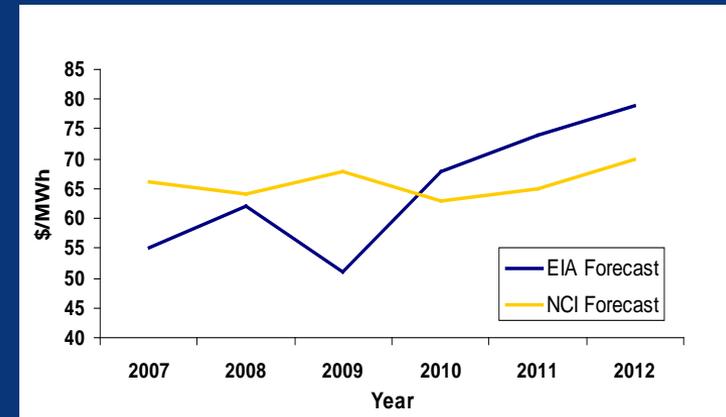
Vlookup Example

SUMIF X ✓ fx =vlookup(B16,B7:C11,2,fasle)

	A	B	C	D	E	F	G	H	I
1									
2									
3									
4									
5									
6									
7		Tariff 1	1						
8		Tariff 2	2						
9		Tariff 3	3						
10		Tariff 4	4						
11		Tariff 5	5						
12									
13									
14									
15									
16		Tariff 3	=vlookup(B16,B7:C11,2,fasle)						
17									
18									
19									
20									
21									
22									

Scenario And Sensitivity Cases

- *What is a scenario?* An alternative view on one or more of the variables in a spreadsheet
- *What is a sensitivity?* Analyzes the impact of changing one or more variables within a probabilistic range
- *Why are scenarios and sensitivities so important in modeling?* Assists with avoiding long, costly model re-designs



Scenario And Sensitivity Case Set Up

- Models without Case architecture have a single input line for each value

Interest Rate	5.5%
Discount Rate	14%
Start Year	2008

- Building models for Case manipulation provides a much more efficient solution

User selects Case to be used – the “Switch”

Case #	2	1	2	3
Case Name	High Case	Base	High Case	Low Case
Interest Rate	5.50%	5.00%	5.00%	5.10%
Discount Rate	12.30%	12.30%	10.90%	13.30%
Start Year	2006	2006	2006	2007

Formulas draw selected case into “active” cells

User entered data (hard-coded)

Case design on Input worksheets depends on whether the input is a Scalar, Vector, or Matrix

Switch

Active Case

Active Case:	1	Test 1				
<u>Coal Prices</u>	Jan-06	Feb-06	Mar-06	Apr-06	May-06	Jun-06
Plant 1	4.5	4.2	5.1	4.9	4.5	4.3
Plant 2	5.1	4.9	5.3	5.2	5.4	5.1
Plant 3	3.9	4.1	4.1	3.8	4.8	4.7

User Cases

	1	Test 1				
<u>Coal Prices</u>	Jan-06	Feb-06	Mar-06	Apr-06	May-06	Jun-06
Plant 1	4.5	4.2	5.1	4.9	4.5	4.3
Plant 2	5.1	4.9	5.3	5.2	5.4	5.1
Plant 3	3.9	4.1	4.1	3.8	4.8	4.7
	2	Test 2				
<u>Coal Prices</u>	Jan-06	Feb-06	Mar-06	Apr-06	May-06	Jun-06
Plant 1	4.2	4.8	5.3	4.7	4.1	4.9
Plant 2	5.3	4.7	5.7	5.3	5.1	5.4
Plant 3	3.8	3.9	3.8	3.9	3.8	3.7

Exercise – Workbook Structure – Input Worksheets

4

What You Need To Do

- Copy the Tariff Model.
- Establish two sets of data in the Billing Determinants Tab and label them “High Case” and “Low Case”.
- Install a switch which instructs the model to use one set of data or the other.
- You have 15 minutes.



Calculation Worksheets Should Include Only Formulas

- Exclude data or “hard-coded” values in worksheets
- Exclude hard-coded values in formulas

For instance, $=A2/(A5+A4)$ is more appropriate than $=A2/(223+423)$

- Break complex formulas into two or more steps in separate rows or columns
- Add Comments (Shift+F2) to cells to explain formulas

Formulas Should Be Consistent

- Develop a formula that can accommodate the required calculation for all cells in the row
- This can usually be accomplished with IF...THEN syntax

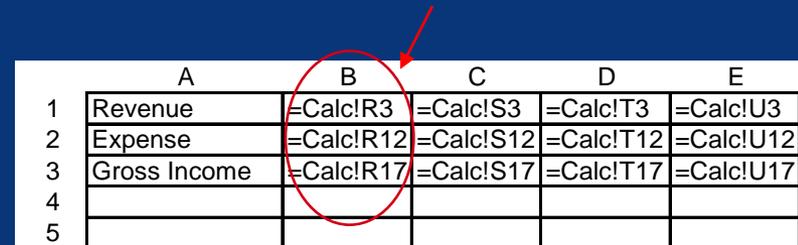
=if(B1=startyr, revenueyr1, A2*(1+EscRate))

	A	B	C	D	E
1		2007	2008	2009	2010
2	Revenue				

Output Worksheets

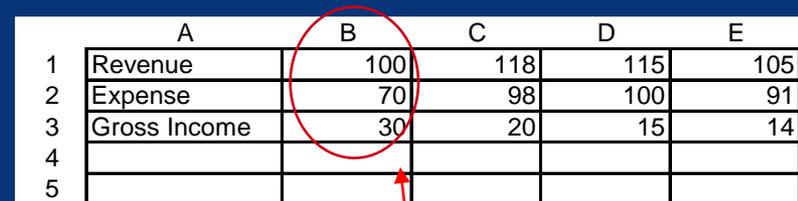
- *Purpose:* To depict the results of the model analysis in a form preferred by the reviewer
- *Types of Output Worksheets:*
 - **Dynamic worksheets** contain all formulas and pull information directly from the calculation worksheets
 - **Static worksheets** are “hard-coded” values that are copied either from calculation worksheets or dynamic worksheets

Dynamic Output worksheets are formula driven



A diagram of a dynamic output worksheet. It shows a table with columns A through E and rows 1 through 5. Row 1: Revenue, =Calc!R3, =Calc!S3, =Calc!T3, =Calc!U3. Row 2: Expense, =Calc!R12, =Calc!S12, =Calc!T12, =Calc!U12. Row 3: Gross Income, =Calc!R17, =Calc!S17, =Calc!T17, =Calc!U17. Row 4: (empty). Row 5: (empty). A red circle highlights the formula in cell B3, and a red arrow points to it from the text above.

	A	B	C	D	E
1	Revenue	=Calc!R3	=Calc!S3	=Calc!T3	=Calc!U3
2	Expense	=Calc!R12	=Calc!S12	=Calc!T12	=Calc!U12
3	Gross Income	=Calc!R17	=Calc!S17	=Calc!T17	=Calc!U17
4					
5					



A diagram of a static output worksheet. It shows a table with columns A through E and rows 1 through 5. Row 1: Revenue, 100, 118, 115, 105. Row 2: Expense, 70, 98, 100, 91. Row 3: Gross Income, 30, 20, 15, 14. Row 4: (empty). Row 5: (empty). A red circle highlights the value 30 in cell B3, and a red arrow points to it from the text below.

	A	B	C	D	E
1	Revenue	100	118	115	105
2	Expense	70	98	100	91
3	Gross Income	30	20	15	14
4					
5					

Static output worksheets are hard-coded versions of Dynamic Output worksheets

Worksheet Arrangement

- Vertical programming
- Data spacing – Don't use empty columns to create space – use formatting

Gross Revenues (\$millions)				
	Company X		Company Y	Company Z
Jan-06	12.5		6.5	17.4
Feb-06	13.5		6.5	22.1
Mar-06	11.8		7.2	23.7
Apr-06	12.2		7.1	21.2
May-06	15.5		8.1	23.9
Jun-06	17.1		9.9	24.7



Gross Revenues (\$millions)				
	Company X	Company Y	Company Z	
Jan-06	12.5	6.5	17.4	
Feb-06	13.5	6.5	22.1	
Mar-06	11.8	7.2	23.7	
Apr-06	12.2	7.1	21.2	
May-06	15.5	8.1	23.9	
Jun-06	17.1	9.9	24.7	

	Jan	Jan	Feb	Feb
	MWh	\$	MWh	\$
Unit 1	100	9,500	98	9,310
Unit 2	112	10,640	101	9,595
Unit 3	140	13,300	122	11,590



(MWh)	Jan	Feb
Unit 1	100	98
Unit 2	112	101
Unit 3	140	122
(\$)	Jan	Feb
Unit 1	9,500	9310
Unit 2	10,640	9595
Unit 3	13,300	11590

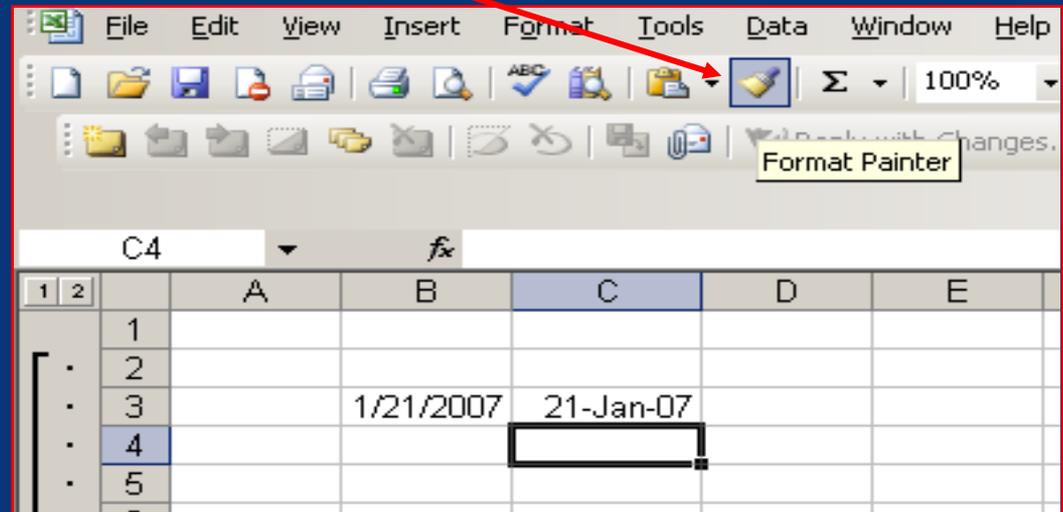
Cell Formatting

- Format values consistently across the spreadsheet
- Column headers generally should be centered, vertical alignment set to bottom, and text wrapped
 - For titles across columns, use Center Across Selection
 - Don't use "Merge cells" – it creates problems
- For presentation, set sheet background color to white and use borders around datasets
- Use text color to designate different types of data, for example:
 - **Blue:** input data
 - **Red:** errors from conditional formatting
 - **Dark yellow:** work in progress, unfinished calculations
 - **Black:** all other



Format Painter

- Format Painter applies the format from one cell or a set of cells to another cell or set of cells
 - **Step 1:** select cell with source format (format you want to apply)
 - **Step 2:** click on Format Painter
 - **Step 3:** select target cell (cell to which you would like to apply source format)
- Alternative: Paste formats only
 - Copy (Edit >> Copy, or Ctrl+C)
 - Edit >> Paste Special >> Formats >> OK



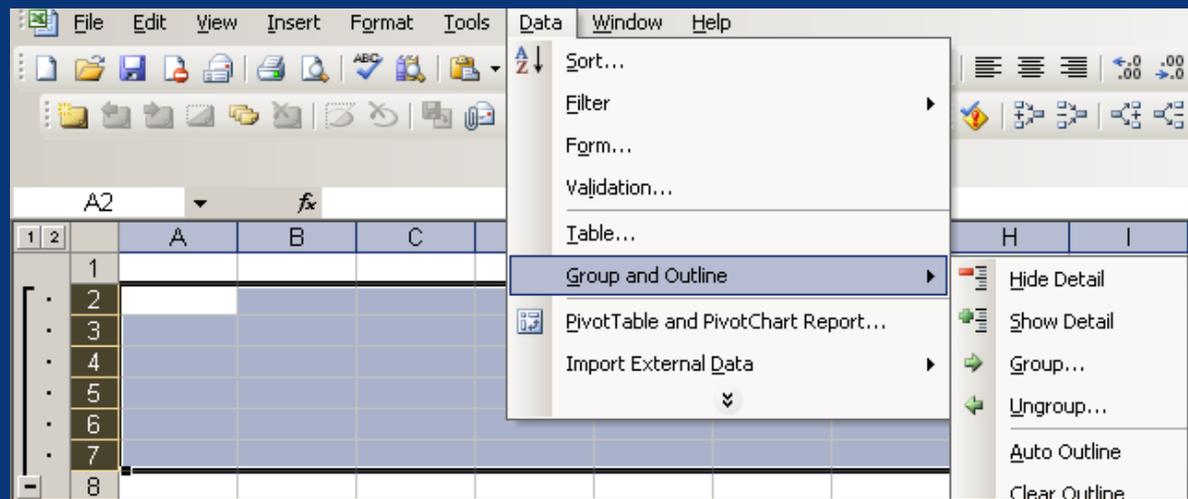
Titles, Headings, Labels and Footnotes

- Titles, headings, and labels should be descriptive and footnotes explanatory.
- Table Title should specify unit where the entire table is in the same units
- Footnotes should be added to output tables to fully explain data presented
- Worksheets should include footers with filename and NCI
- Comments should be added to input cells to explain calculations and data sources (Shift+F2)
- Thousands and millions can be represented in the following ways:
 - Thousands: “(thousands),” “(000),” (“\$000”)
 - Millions: “(millions),” “(000,000),” (“\$000,000”)

Column And Row Size

- Column width should be consistent when displaying data of a similar nature
- Row width generally should be subordinate to column width
- Rows and Columns can be hidden
- Grouping allows users to collapse data to improve spreadsheet flow and organizing calculations

Grouped rows



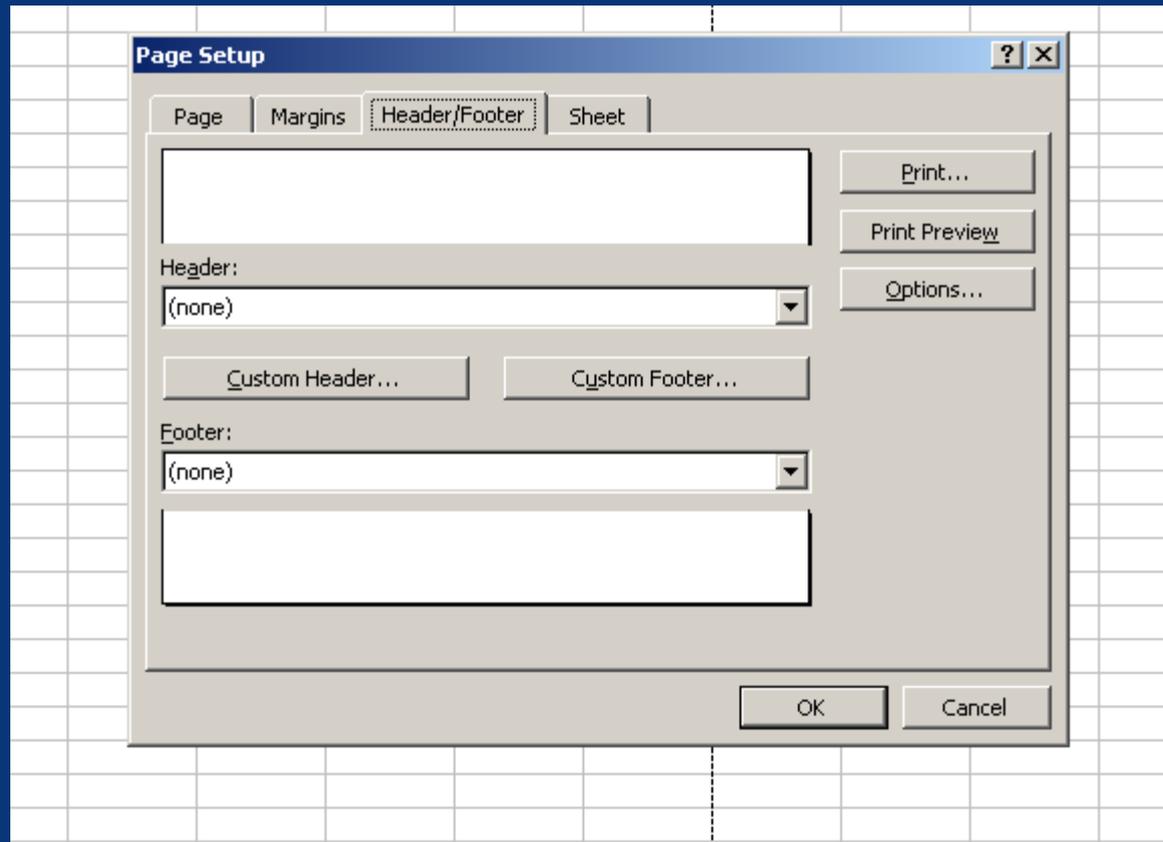
Naming Worksheets

- Worksheets should be named consistently with tab color and 'hiding' used effectively to improve the interface
- Individual worksheet type (data, calculation, output, switch) should be designated either in the worksheet name or by tab color
 - **Light blue:** input data and switches
 - **Light yellow:** calculations
 - **Light green:** output
- Worksheets can be hidden using Format>Sheets>Hide

Print Formatting

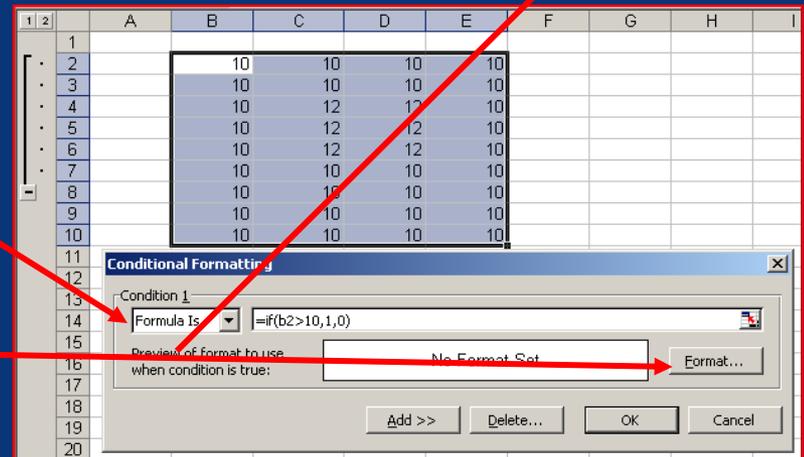
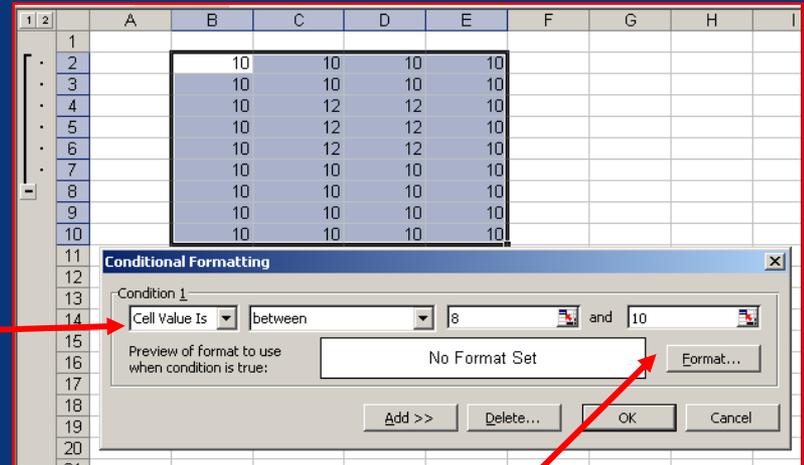
- Do not provide worksheets that are not pre-formatted for printing
- File>Page Setup
 - Page
 - Select portrait or landscape depending upon size of spreadsheet
 - Select “Fit to” and either designate the number of pages wide OR the number of pages tall that will present your data in a readable format.
 - Review impact of formatting by clicking Print Preview
 - Headers/Footers
 - Can select header/footer from drop down or add custom header/footer
 - Should always include a footer on any material likely to be printed and/or distributed
 - Sheet
 - Specify worksheet print area (default is all sections where values entered)
 - Specify whether certain columns or rows should repeat on printed sheets

Inserting a Header or Footer



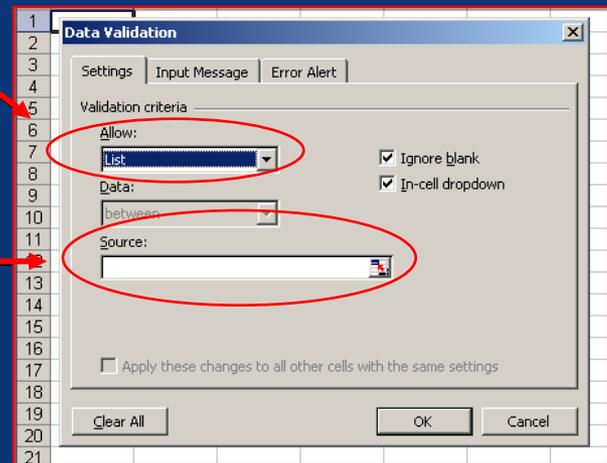
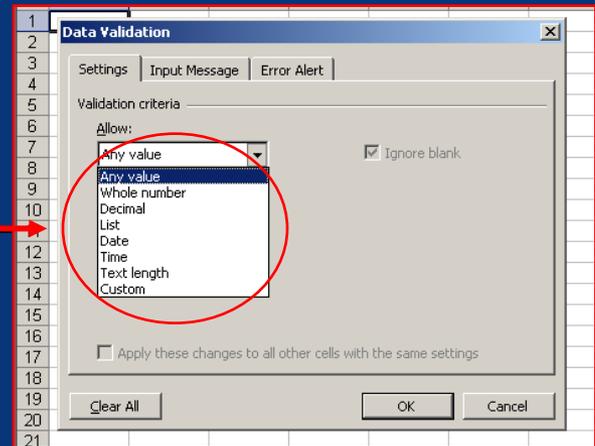
Conditional Formatting

- Cell value is...
 - Specify the cell is $<$, $>$, $=$ (etc) than particular values (either hand-entered or reference to cells)
- Formula is...
 - Enter a formula that resolves to a True or a False.
 - For any multi cell calculations across the selected cells, the relative reference begins with the cell in the left upper-hand corner
- Then click Format button and choose the style to apply to the cells that meet the criteria
- Limitation: If used extensively, slows calculation



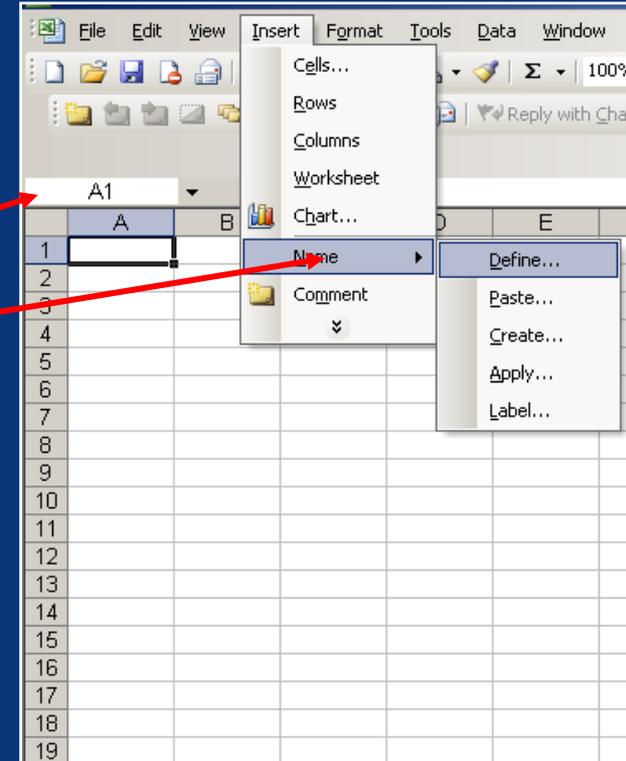
Data Validation

- Data>Validation
- To limit value based on criteria
 - Data>Validation>“whole number,” “decimal,” “date”
 - Enter min/max to limit value of cell
- To limit value based on defined set of values
 - Data>Validation> “List”
 - In the Source box, one can either (1) manually enter list, (2) refer to specific cells (e.g., =C3:G3), (3) refer to a range name (e.g., =testlist)*
 - Cell with data validation will now have a drop down box



Range Names

- *Purpose:* Greatly reduces errors, improves readability and increases modeling productivity
- *Assigning Range Names (Two options):*
 - Highlight cells and type into Name Box then press Enter
 - Insert>Name>Define and type formula
- *Defining Range Names:* Use a formula or a straight cell reference
- *Applying Range Names:* Apply to all model constants
- *Documenting Range Names:* List and describe range names in the documentation sheet



Worksheet And Workbook Protection

- *Purpose:* Workbook and Worksheet Protection are important to ensuring the integrity of the finished model
- *Protecting Workbooks:* Tools>Protection>Protect Workbook
 - Protects the structure of a workbook
 - Protects windows so that they are the same size and position each time the workbook is opened
- *Protecting Worksheets:* Tools>Protection>Protect Worksheet
 - Protects a specific worksheet from changes to cells
 - Will protect the entire sheet unless modeler specifies otherwise in advance

Documenting Worksheet Programming Changes

Programming documentation can be accomplished using two methods:

- Separate worksheet to log changes: preferred method to track intra-version changes

	Name Box	B	C	D	E	F	G	H
1								
2								
3	Model Changes							
4	10/14/2006	Fixed depreciation table on Calculations tab for first year issue						
5	10/25/2006	Developed interest rate calculation for different tranches of debt from corporation						
6	11/5/2006	Updated data set with company third quarter filings						
7								
8								
9								

Questions



Table of Contents

1	Introduction
2	The Model Lifecycle
3	Model Scoping and Specification
4	Model Design
5	Model Construction
6	Model Documentation
7	Model Testing

Objectives

- Construct dynamic spreadsheet models in order to achieve flexibility, efficiency and accuracy
- Efficiently manipulate text, dates and numerals
- Apply advanced Excel functions that aid in identifying and selecting data
- Construct a model related to the case study for this class



Data Manipulation: Dates

Date Functions

=date(year,month,day)

- Turns a year, month, and a day into full date
- DATE(2007,13,1) return Jan. 1, 2008
 - useful for dealing with the next month, even if it's in another year

=month(date)

- Returns month number

=year(date)

- Returns year number (1900 – 9999)

=weekday(date)

- Returns day of week (Sunday = 1)
- =weekday(date,2) makes Sunday = 7

=edate(start_date,X)

- Returns date X number of months from start date
- Requires Analysis ToolPak

=eomonth(date,X)

- Returns last day of month that is X months from date

=networkdays(start_date,end_date,holidays)

- Returns number of work days adjusted for holidays
- Ignores holidays unless a list is specified

Data Manipulation: Numbers

=Sumif(range, criteria, sum range)

- Range is range to be evaluated
- Criteria is the condition the range is evaluated against
- Sum range is range to be summed if condition is true

=Sumproduct(array1, array2,...) or =Sumproduct(array1*array2*...)

- multiplies corresponding components in the given arrays, and returns the sum of those products
- From 2-30 arrays may be used

=Sumproduct(array1*array2*(array3=value)*...)

- multiplies components if conditions are met
- e.g., add up price x quantity for all hours in January

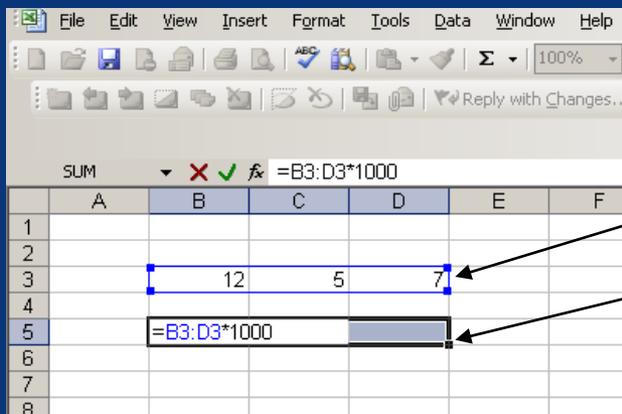
Data Manipulation: Arrays

- An *array* is simply a string or matrix of data: numbers, text, etc.
- An *array formula* can perform multiple calculations across an array and then return either a single result or multiple results
 - *Multiple Result Array Formula*: A formula entered into a group of highlighted cells that performs an operation across an equivalent sized set of cells, returning a result for each of the highlighted cells.
 - *Single Result Array Formula*: A formula entered into a single cell that evaluates each cell in a target array of cells against a specified criteria then summarizes those cells meeting the criteria using a particular summary operation (e.g., sum, average, etc).
 - Single or Multiple - must be entered with Ctrl+Shift+Enter

Data Manipulation: Arrays (Continued)

Multiple result array formulas:

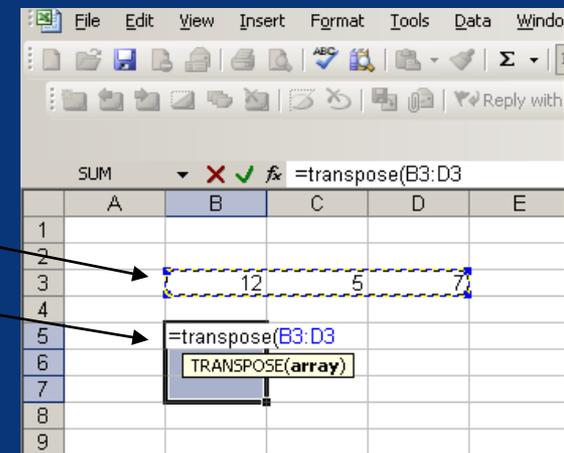
- Can be used for transposing data and performing operations across arrays.
- Formulas are entered into multiple cells simultaneously and perform a calculation across single or multiple arrays of data



Existing Array

New Array

Note, range of cells highlighted and single formula entered.



Data Manipulation: Array Formulas

Single result array formulas:

- May be the most useful function for spreadsheet modeling—avoids “in-between” calculations and macros
- Formulas perform calculation across a source array and return a single value
- Most common type: Determine average, sum or other based on multiple conditions



Data Manipulation: Numerals - Arrays

Single result array formulas:

Example: find the average revenue of NY companies with employees greater than 80

`=AVERAGE(IF(C3:C14="NY",IF(D3:D14>80,E3:E14)))`

Summary Function

Condition 1

Condition 2

Argument

	A	B	C	D	E	F
1						
2		Company	State	Employees	Revenue (\$M)	Net Income (\$M)
3		Company A	MA	120	50	4.3
4		Company B	NH	99	43	2.1
5		Company C	MA	55	70	7.2
6	→	Company D	NY	170	125	5.4
7		Company E	RI	72	44	1.2
8	→	Company F	NY	75	92	0.8
9		Company G	RI	78	140	0.4
10		Company H	MA	81	188	1.8
11	→	Company I	NY	84	110	4.1
12	→	Company J	NY	87	76	17.2
13		Company K	NH	90	54	12.1
14	→	Company L	NY	93	122	1.1
15						
16						
17						108.25

The same result can almost always be achieved without an array formulas using Sumproduct (with * not ,):

`=SUMPRODUCT((C3:C14="NY")*(D3:D14>80)*E3:E14)/SUMPRODUCT((C3:C14="NY")*(D3:D14>80))`

Data Manipulation: Identifying & Selecting Data

Don't use VLookup and HLookup Functions:

- VLookup: “=vlookup(table,value,#columns to right)”
- HLookup: “=hlookup(table,value,row#)”
- Can be unstable and inefficient functions leading to errors:
 - Column and row offsets MUST remain constant
 - Columns and rows cannot be offset to left or up
 - Lookups cannot be based on multiple criteria

Index>Match can be used instead in all cases

Data Manipulation: Identifying & Selecting Data

Index/Match Functions:

- Uses two arrays to find the target value:
 - Lookup Array – Similar to leftmost column in VLookup
 - Target Array – Similar to column designated in VLookup
- Syntax: “=index(target array, match(lookup value, lookup array,0))”
- Robust tool for identifying and selecting data
 - Eliminates need for VLookup

Data Manipulation: Identifying & Selecting Data

Index/Match Example:

Example: find the net income of Company D

`=INDEX(F3:F14,MATCH(B19,B3:B14,0))`

	A	B	C	D	E	F
1						
2		Company	State	Employees	Revenue (\$M)	Net Income (\$M)
3		Company A	MA	120	50	4.3
4		Company B	NH	99	43	2.1
5		Company C	MA	55	70	7.2
6		Company D	NY	170	125	5.4
7		Company E	RI	72	44	1.2
8		Company F	NY	75	92	0.8
9		Company G	RI	78	140	0.4
10		Company H	MA	81	188	1.8
11		Company I	NY	84	110	4.1
12		Company J	NY	87	76	17.2
13		Company K	NH	90	54	12.1
14		Company L	NY	93	122	1.1
15						
16						
17						
18						
19		Company D	5.4			

Data Manipulation: Identifying & Selecting Data

Offset Function:

Syntax: =OFFSET(reference,row,column,height,width)

- Can return a single cell or an array for further calculation

Example: Calculate the average hours for first 3 months of the year
=AVERAGE(OFFSET(A2,0,0,1,3))

- Formula takes average of cells A2 through C2

	A	B	C	D	E	F
1	Jan-07	Feb-07	Mar-07	Apr-07	May-07	Jun-07
2	4.5	4.2	5.1	4.9	4.5	4.3

Data Manipulation: Identifying & Selecting Data

Indirect Function:

- **Syntax: =INDIRECT(reference)**
 - Evaluates contents of the cell referenced by the 'reference cell'
 - Commonly references a range name
- **Very useful in multi-sheet programming to avoid formula differences between sheets**
 - If two worksheets contain the same calculations but rely on different data sources, Indirect can be used to pull correct data
 - Can be used to pull similar data from the same place in many sheets at the same time

Data Tables

- **Runs through a series of values for one or two variables and records the results**
 - Used for sensitivity cases, or for complex calculation that applies to
- **For a single variable:**
 - Set up the control cell: the switch or key value that controls the calculations
 - To the right, create formulas representing the values to record
 - Below, list the alternative values
 - Data >> Table >> Column input cell:
- **For two variables:**
 - In one cell, create a formula representing the value to record (one formula only)
 - Below, list alternative values for one variable
 - To the right, list alternative values for the other variable
 - Set up two cells on the same sheet to be the control variables
 - Data >> Table >> Row input cell AND Column input cell
- **Usage notes**
 - Great for running through all sensitivity cases in a single table
 - Useful for applying the same calculations to several instances
 - In a complex spreadsheet, a data table can slow calculation to a crawl.
 - Use Tools >> Options >> Calculation >> Automatic except tables

Model Design Exercise

What You Need To Do

- Clean up the tariff model. Put the sheets in a logical order. Color-code both the sheets and ranges within the sheets.
- Experiment with some of the techniques and functions you have just learned.
- Complete by tomorrow morning.

Data Manipulation: Questions



Table of Contents

1	Introduction
2	The Model Lifecycle
3	Model Scoping and Specification
4	Model Design
5	Model Construction
6	Model Documentation
7	Model Testing

Documentation Objectives



- Understand the importance of documenting assumptions and instructions in models
- Recognize best practices for writing effective documentation within any project model
- Take the documentation skills learned and apply them to any work project

Purpose of Documentation

- Enable others to use model and understand its purpose, long after it was created and without the original creator
- Reduce risk of misinterpretation of findings
- Improve efficiency in review



Best Practices Discussion



- Describe some best practices you use to document assumptions and criteria in your models
- What value have you found in these practices?
- Provide an example of a situation in which best practices for documentation were not followed
- Describe the results

Best Practices Model Documentation



- Provide purpose description
- Identify modeler, contact information
- Define model structure/file layout
- Summarize model
- Delineate switches, data, inputs, calculations and reports
- Provide instructions, assumptions

Best Practices *(Continued)*

Worksheet Documentation

- Use appropriate headers/footers
- Identify engagement on every sheet
- Identify worksheet purpose in title
- Identify date ranges for data/analyses
- Use range names



Best Practices *(Continued)*

Referencing/Source Documents



	+2.688
	+5.000
	+1.500
	+1.125
	+1.062

- Distinguish notes/references from analysis
- Use a consistent reference method
- Consistently place references
- Source all hard-coded data
- Source all assumptions
- Explain formulas where not obvious

Best Practices

(Continued)



- Worksheet Labeling
 - Designate individual worksheet type in the worksheet name or by tab color
 - **Light blue:** input data and switches
 - **Light yellow:** calculations
 - **Light green:** output
 - Keep spreadsheet design understandable to the reviewer
 - Hide worksheets using Format>Sheets>Hide, which is a more acceptable practice than hiding rows and columns when the number of sheets becomes unmanageable

Best Practices

(Continued)

- **Worksheet Documentation**
 - Make all titles, headings, labels explanatory
 - Specify units in the Table Title when the entire table is in the same units
 - Add footnotes to output tables to fully explain data presented, including data source and any nuances
 - Include footers with filename and NCI name
 - Add comments to input cells to explain data sources and references
 - Consistently represent thousands and millions



Best Practices

(Continued)



- Documenting Workbook Changes
 - Document changes made to a model to avoid missing “intended” edits and provide an explanatory tool
 - Create a separate worksheet to log intra-version changes made in the workbook
 - Consider using track changes mode to clearly define changes made and not made

Questions?

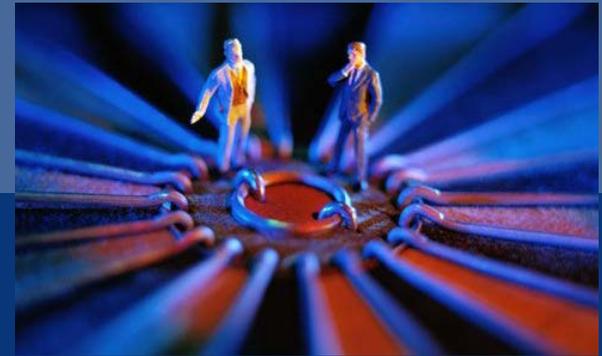


Table of Contents

1	Introduction
2	The Model Lifecycle
3	Model Scoping and Specification
4	Model Design
5	Model Construction
6	Model Documentation
7	Model Testing



Testing Objectives



- Understand the importance of testing spreadsheet models
- Be aware of error types to identify
- Employ a rigorous approach to testing models
- Employ several testing techniques and apply them appropriately to the model's structure and circumstances
- Both self-test models and test the models of others, as well as provide feedback to others on issues to be addressed

Consequences Of Poor Modeling

Modeling error consequences:

- ... One strategic alternative looks more attractive than another when the opposite is true, costing a company millions
- ... An expert witness is subject to intense cross-examination, is excluded as a witness
- ... A public company is forced to restate earnings because the model supported financial reporting



Keys To Remember

- Start testing with the assumption that something is wrong in the model
- When one does not think one needs to quality control, that is when quality control is most needed
- While self-review is important, it is also understandable that one can be too close to one's own model—independent review is also necessary



Reasons For Errors

- Models are developed iteratively
- Even the most careful person will miss a change in a model, make an error in referencing or formula, forget to make a change in logic, etc.
- The better the design of the model, the easier it is to identify and fix errors in testing—as well as the easier the testing process

Types Of Errors

- Mechanical errors
 - Wrong formula or value is entered in a cell
- Logical errors
 - Misunderstanding of the problem being addressed or a misapplication of theory in solving the problem
- Omission errors
 - Model inappropriately excludes a piece of data or a calculation



Examples Of Common Errors

- Formula not copied correctly
- Formula updates or changes not carried forward or copied in all places the formulas are used
- Wrong cell reference
- Using a semi-absolute reference when one wanted to use an absolute reference and vice versa (e.g., \$D4 versus \$D\$4)
- Summation does not include all rows or columns intended in a range, particularly when a row or column was added to the beginning or end of a range



Examples Of Common Errors

(Continued)

- Including a subtotal in a summation; double-count
- Using the wrong unit of measure (such as 1000's versus 1,000,000's, or 0.1 versus 0.001)
- Using a function inaccurately—such as VLOOKUP, HLOOKUP, MATCH
- Error in a complex formula—note all the common errors one can make without involving a complex formula



Testing Techniques, Uses And Limitations

- What are some testing techniques?
- What are some uses of each?
- What are some limitations of testing techniques?



Testing Techniques, Uses And Limitations *(Continued)*

- Eyeballing
- Signing off
- Auditing
- Execution Testing
- Static Testing Programs
- Cell by cell review



Beginning Model Testing – The Logic Review



- What is the purpose of the model and does it achieve it?
 - What financial theories or relationships are involved?
 - Are there commonly understood relationships or ratios that should be employed?

Beginning Model Testing – The Logic Review *(Continued)*

- How does the model flow?
 - Were Bubble Charts, Variable Outlines or Calculation Tables used?
 - If so, have they been updated to reflect all the iterative changes to the model?
 - Are the logical relationships still intact?



Model Testing: Cell-By-Cell

- Uses Excel's audit tools
- In general, calculation cells within the same row will have the same formula
 - This can be verified using Excel's Error Checking (possible errors are shown in top left corner) or manually
 - Thus, only the first cell's formula will need to be evaluated



Model Testing: Cell-By-Cell (Continued)

- Separate model inputs from calculations
- Avoid constants within calculation cells
- Begin with the model output and proceed backward toward inputs using Excel's audit tools, such as Tracing



Model Testing: Cell-By-Cell (Continued)

- Diagram the calculation and work backward through each tree individually
- Use “Evaluate Formula” where formulas become difficult to dissect or produce aberrant results
- At least two people should perform a detailed review of any model of significance



When To Test And Quality Control Models

- Model complete; changes not envisioned
- Before the model is produced to the client or an expert report is filed
- After changes have been made to a model that has already been tested



Documentation Of Review

- For the detailed review of the model, the reviewer should print a hard copy of the model
- Use tickmarks in each cell to indicate the review
- Include a tickmark legend to indicate the meaning of each tickmark, which documents the work that he/she has done
- Insert a tab within the model spreadsheet
- On this tab, create a matrix with the following headings, and document comments

Worksheet Tab

Cell Reference

Comments



Model Testing Exercise

What You Need To Do

- Compare the results of your revised spreadsheet with the original spreadsheet.
- Document any differences in an e-mail to the instructors. Include recommendations.
- This exercise is assigned as homework.

Questions

