Rich Malloy, NCC Extended Studies

Session 3:
Advanced Excel 20116

ANALYZING BIG DATA

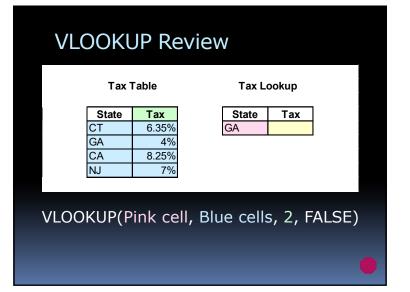## How Big Can a Worksheet Get?

- Over 16,000 columns

- Over 1 million rows

- But, over about 500,000 rows:
  - File is very large
  - Program is slow
  - Program can be unsteady

Big Data

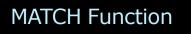- As you and your company grow

- Excel can grow with you

Agenda

- Review VLOOKUP

- Beyond VLOOKUP: INDEX/MATCH

- Workbook Management & Design

- Databases: Flat File vs Relational

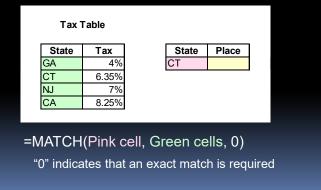- The Excel Data Model

- Power Pivot

VLOOKUP Review

Tax Table

State	Tax
CT	6.35%
GA	4%
CA	8.25%
NJ	7%

Tax Lookup

State	Tax
GA	

VLOOKUP(Pink cell, Blue cells, 2, FALSE)

VLOOKUP Problems

- Lookup column has to be the first column in the lookup table
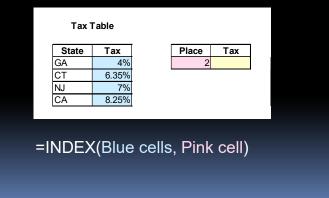- Cannot copy across columns
- Cannot do a 2-dimensional lookup

MATCH Function

- **Finds** location of a particular value

Tax Table

State	Tax
GA	4%
CT	6.35%
NJ	7%
CA	8.25%

State	Place
CT	

=MATCH(Pink cell, Green cells, 0)

"0" indicates that an exact match is required

INDEX Function

- **Gets** a value from a particular location

Tax Table

State	Tax
GA	4%
CT	6.35%
NJ	7%
CA	8.25%

Place	Tax
2	

=INDEX(Blue cells, Pink cell)

INDEX & MATCH

- Finds location of a particular value

- And gets it

- Can replace VLOOKUP

INDEX & MATCH

Tax Table			Tax Lookup	
State	**Tax**		**State**	**Tax**
CT	6.35%		GA	
GA	4%			
CA	8.25%			
NJ	7%			

=INDEX(Blue cells, MATCH(Pink cell, Green cells, 0))

INDEX/MATCH Advantages

- Can lookup values in 2nd or 3rd column

- Can be copied more easily

- Can return values from a row

- Can do a two-way lookup

Power Pivot

- Formerly an extra-cost option

- Now a free add-in for Excel 2016

- Runs on the Power Pivot Engine
  - Also known as the **Data Model**

The Data Model: Advantages

- Millions of rows of data

- Relationships between tables

- "Easier, faster than VLOOKUP"

- Powerful new functions
  - E.g., the CALCULATE function
  - Unique Counts

Related Tables

- Tables linked by common values

- Part of the Data Model
  - Also key part of Relational Databases

- Uses less storage space

- Very efficient and flexible

- See "Parents & Children Slides"

The Data Model: 2 Approaches:

- Excel

- Power Pivot

Create a Data Model in Excel:

1. Convert data into Excel Tables
   - Use: Format as Table tool

2. Use one table to create a Pivot Table
   - Click: Add to Data Model

3. Field List: All (instead of Active)

4. Choose a field from each table

5. Create a relationship: Auto-Detect

The DAX Language

- DAX: Data Analysis Expressions

- DISTINCTCOUNT function
  - Counts Unique values

- CALCULATE function
  - "SUMIF on steroids"
  - Powerful filtering options

PowerPivot Advantages

- Create table relationships visually

- DAX calculations

- Can define Hierarchies

- Can create Key Performance Indicators

- Can import data from multiple sources

PowerPivot Disadvantages

- Cannot edit individual cell data

- Cannot use VBA to automate tasks

PowerPivot Resources

- Mike Girvin
  - PowerPivot Intro
    - https://www.youtube.com/watch?v=yjASjwJCIxA

REVIEW

Which Function will add up numbers linked to a particular category?

1. SUM
2. SUMIF
3. SUMIFS
4. SUBTOTAL

Which Function will add up numbers linked to a particular category?

1. SUM
2. **SUMIF**
3. **SUMIFS**
4. SUBTOTAL

SUMMARY

Use Excel Tables

Category	Jan	Feb	Mar	Apr	May	Jun	Total	% of Total
Rent	2,002	2,004	2,006	2,008	2,010	2,012	12,042	48.4%
Auto Loan	300	300	300	300	300	300	1,800	7.2%
Utilities	250	250	250	200	150	150	1,250	5.0%
Groceries	900	900	900	1,100	900	900	5,600	22.5%
TV	152	154	156	158	160	162	942	3.8%
Savings	202	202	202	202	202	230	1,239	5.0%
Misc	250	250	400	500	200	400	2,000	8.0%
Total	**4,056**	**4,060**	**4,214**	**4,468**	**3,922**	**4,154**	**24,873**	**100.0%**

Use Slicers to Filter Data

Category

Auto ...	Groce...	Misc
Rent	Savings	TV
Utilities		

Month

Jan
Feb
Mar