Intermediate Excel: Reference Functions and Pivot Tables

Understanding the most important analytical tools in excel

Disclaimer

The purpose of this tutorial is simply to provide an introduction to a variety of topics in Microsoft Excel. It is in no way a substitute for a full for-credit course on Excel, and should students want to pursue Excel on a more serious level they should look to the university courses on the subject

Before starting the tutorial: make sure you download the Excel workbook with the sample dataset. It should contain about 200 rows and 12 columns with various pieces of information about movies from IMDB.

Reference Functions
VLOOKUP()

These value look up formulas are one of the most intimidating excel formulas, but once you get the hang of them, they are quite simple. Here is the basic form:

=VLOOKUP(lookup_value, table_array, col_index_number, [range_lookup])

Lookup_value: The value that we want to find

Table_array: The place that Excel will look for your data
*Note: The LEFTMOST column in your array will always be read as your *index ,* this is where Excel will go to find the value you are looking up

Col_index_number: The number of the column (in your array) containing the values that you want returned back to you

[range_lookup]: Decide if you want an Exact Match or Approximate Match – indicated as FALSE or TRUE

Practice Problem:

What Genre is the movie with id=45?

=VLOOKUP(Movie ID we are looking for (in this case, 45),Table Range for the Movies Database, Column number for Genre, TRUE)

You should get:

 Action|Adventure|Sci-Fi

We can alternatively make our own unique ID column using a CONCAT function, depending on what our ultimate goal is

Difficulties of the VLOOKUP()
One of the issues with the VLOOKUP function is the requirement of creating an index, as well as the rigid nature of defining our search column. The combination of using the functions index and match help solve this.

INDEX(), MATCH()
Index and match are two functions that when used in conjunction with one another provide an alternative to Vlookups that is in many ways more dynamic. Where a in a vlookup, we needed some sort of unique index, using index match allows us to index the value we want simply based on the column and row titles.

Index = (array, row_num, [column_num])
This should simple look in the preselected array and retrieve the value that is in whichever row and column you have specified.

Array: The place that Excel will look for your data

Row_num: The row in which Excel will look for your data (this is a number)

[column_num]: The column in which Excel will look for your data (also a number)

Example:
Find the value in the 100th row and 5th column of our dataset

Match = (lookup_value, lookup_array, matchtype)
Match is like the inverse of the index function. You start by specifying a value you want to find, then you look in a specific array for that value. The output will be the location of the data.

Lookup_value: Value you want Excel to find

Lookup_array: The place that Excel will look for your data (Make sure you pick an array with only one dimension of data-a single column or row)

Matchtype: The type of match that you want

Example:
Find the row location of the film that has Andrew Adamson as the director

Together:
INDEX(total array, MATCH(for row value), MATCH(for column value))

Pivot Tables
Pivot tables are traditionally considered the most powerful tool that is offered by microsoft excel. They are to date the fastest way to quickly rearrange, aggregate and analyze datasets in

Excel. While pivot tables originated in the mid 90's, and were then joined by pivot charts in the early 2000's, many people are still unfamiliar with them.

Lets Create a Pivot Table
- Go to the **Insert** Tab, select Pivot Table, and select our movies database as our range

Field List
- Fields
- Values
- Rows
- Columns
- Filters
- Value Field Settings

Pivot Table Tabs

Analyze
- Slicer/Timelines
- Refresh Data

Design
- Totals/Report Layout

Pivot Charts (Not available on Macs)

Practice Exercises:
1. Does the release year seem correlate with the number of facebook likes?
2. Do Certain Genre's do better on average in the box office than others?
3. What Percentage of the movies were rated PG-13? PG?