Rev Up Your Data Skills
Data Clean-up Tips & Tricks

VLOOKUP or INDEX MATCH

VLOOKUP looks for a value in a column and once it finds a match it will return a value from any cell in the same row as the match. It requires a common cell, such as an Employee ID. It can only find information to the right of the common cell.

INDEX MATCH is a combination of two functions which yields the same result as VLOOKUP with the benefit that it can return data from columns to the left or right of the matching value column and doesn't depend on the order of the columns when used with table formatting.

VLOOKUP or INDEX MATCH are great tools for a variety of data checks:

* Are all prior year employees accounted for on the current year's Employee Roster, Terminations Roster or the Promotions/Transfers Rosters?
* Does everyone on the current year's Termination Roster with a current year hire date appear on the New Hires Roster?
* Do any of the Terminations appear on the Employee Roster?
* Are there any job title changes on the Employee Roster or Terminations Roster from the prior year job title?
* Do all of the New Hires appear on either the Employee Roster, Terminations Roster the Promotions/Transfers Rosters?
* Do all of the New Hires appear on the Applicant Flow Log?
* Is there hiring action on all Req IDs that appear on the Applicant Flow Log?

VLOOKUP

=VLOOKUP(Step1_Common_Cell,Step2_Lookup_Range,Step3_Result_Column_Number,0)

=VLOOKUP(A4,'PY Employee Roster'!$A$4:$E$142,5,0)

=VLOOKUP(	The function
A4,	Common cell in the starting file *(in example Employee ID)*
'PY Employee Roster'!	File/spreadsheet name where the data we want is located
$A$4:$E$142,	The range of cells to look for information
5,	Number of columns to the right of the common cell where the data we want is located
0)	Provide only information where the common cell matches exactly

INDEX MATCH

=INDEX(Step3_Result_Column,MATCH(Step1_Common_Field,Step2_Lookup_Column,0))

=INDEX(HiresRoster[Hire Date],MATCH([Employee ID],HiresRoster[Employee ID],0))

=INDEX(	The function
HiresRoster[HireDate],	Table/Column name of the data we want to retrieve
MATCH(	The second function to compare the fields
[Employee ID],	Common column name in starting file *(in example Employee ID)*
HiresRoster[Employee ID],	Common column name from table we are pulling information
0))	Provide only information where the common column name matches exactly

Jennifer McAlister, Director of Operations
Jen@KairosServicesInc.com
Main 972-369-0015
Direct 972-369-0104
Mobile 940-390-7069

KSI
Kairos Services Inc.
Affirmative Action & EEO....our only specialty
2019 ILG National Conference
Milwaukee, Wisconsin

Will Waymel, Director of Audit Services
Will@KairosServicesInc.com
Main 972-369-0015
Direct 972-529-9495
Mobile 214-548-2558

Pivot Tables

Pivot Tables allow you to reorganize and summarize selected columns and rows to obtain a desired report.

Pivot Tables are a great tool for even more in-depth data checks:
* Does every job title have only one job group and one census code?
* Does every department have a managerial employee (EEO-1 Category 1)?
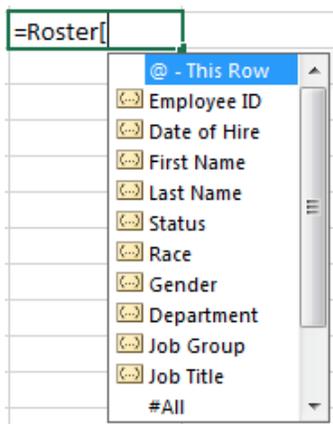* Do the census codes within a job group make sense?

Table Formatting

Converting a data range into a table enables additional functionality in Excel for more efficiency and effectiveness. By using the table features, you can then manage the data in the table rows and columns independently from the data in other rows and columns on the worksheet.

Benefits of table formatting:
* When using formulas you can see and reference table and column names instead of fixed or relative cell references in formulas.
    Example   =A2 * B2   *becomes*   =[Hours Worked] * [Hourly Rate]
* The table expands with additional rows/columns (no need to re-define areas in pivot tables, fill formulas down).
* Unable to disjoint the data by improperly unfiltering/refiltering.
* Formulas fill down automatically to all cells in column every time you make changes.
* When referring to a table on another sheet there is no need to go to the other sheet and look for column names, order, count columns, etc.

How to use Formulas in Tables

=EmployeeRoster[Employee ID]

| =Roster[|
| @ - This Row |
| Employee ID |
| Date of Hire |
| First Name |
| Last Name |
| Status |
| Race |
| Gender |
| Department |
| Job Group |
| Job Title |
| #All |

* A table is referenced by its name - the above is called "EmployeeRoster" (the default is "Table1" if you don't rename it.
* A column (or cell in a column) is referenced by the column header name inside brackets - the above is [Employee ID].
    Note: If you click on a single cell in a table, the formula Excel automatically populates uses an @ symbol like this [@[column name]] but you don't have to if you're typing it out. Excel can figure out if you only want to refer to the single cell in the row or the entire column based on the rest of your formula.
* When writing a formula you can begin typing either a table name and hit tab to fill in the rest or type a beginning bracket "[" and get a list of all columns in that table that filters down to matching column headers as you begin typing.
    Note: When you are referencing cells inside the same table as your formula you don't need to refer to the table name.

Power Query "Get & Transform"

Power Query is an Excel add-in that extracts data, transforms it and loads the data into Excel.

Power Query, now called Get & Transform, can take data skills to the next level:
* Clean data (remove unwanted rows, columns, replace values, etc.)
* Reshape data (convert unstructured or "messy" data into a normalized database format)
* Combine data (12 monthly reports, 100 separate worksheets, etc.)
* Convert/combine multiple files at once (Text, CSV, Excel)
* Perform simple or complex transformations (split columns, unpivot columns, filter data, sort data, etc.)
* Refresh/recreate steps taken with new data in one click (re-use daily, monthly, annually, etc.)