

**This content was specifically designed for use and free distribution on**

***AuditSoftware.Net***

*A community of professionals helping you to maximize the benefits of audit software*

# Using Excel as an Audit Software

**By Richard B. Lanza, CPA, PMP**

# TABLE OF CONTENTS

|                                                  |           |
|--------------------------------------------------|-----------|
| <b>COPYRIGHT PAGE</b>                            | <b>3</b>  |
| <b>DOCUMENT PURPOSE / CALL FOR FEEDBACK</b>      | <b>4</b>  |
| <b>LIMITATIONS OF EXCEL AS AN AUDIT SOFTWARE</b> | <b>5</b>  |
| <b>AUDIT SOFTWARE TESTS</b>                      | <b>6</b>  |
| <b>1. Analytical Tests</b>                       | <b>6</b>  |
| 1.1 Horizontal Analysis                          | 6         |
| 1.2 Vertical Analysis                            | 7         |
| 1.3 Ratios                                       | 7         |
| 1.4 Trend analysis                               | 8         |
| 1.5 Performance Measures                         | 8         |
| 1.6 Statistics                                   | 8         |
| 1.7 Stratifications                              | 10        |
| 1.8 Aging                                        | 11        |
| 1.9 Benford's Law                                | 11        |
| 1.10 Regression                                  | 12        |
| 1.11 Monte Carlo simulation                      | 13        |
| <b>2. Data Management/Analysis Reports</b>       | <b>14</b> |
| 2.1 Append/Merge                                 | 15        |
| 2.2 Calculated Fields/Functions                  | 16        |
| 2.3 Cross Tabulate                               | 17        |
| 2.4 Duplicates                                   | 20        |
| 2.5 Extract/Filter                               | 21        |
| 2.6 Export                                       | 23        |
| 2.7 Gaps                                         | 23        |
| 2.8 Index / Sort                                 | 23        |
| 2.9 Join / Relate                                | 24        |
| 2.11 Summarize                                   | 28        |
| <b>ABOUT THE AUTHOR</b>                          | <b>29</b> |

# Copyright Page

© 2003 - Richard B. Lanza

No part of this publication may be reproduced in any form without permission in writing from Richard B. Lanza.

## *Limitation of Liability / Disclaimer of Warranty*

The author, Richard B. Lanza, has used his best efforts in preparing this publication and is not responsible for any errors or omissions. He makes no representations or warranties with respect to the accuracy or completeness of the contents of this document and specifically disclaim any implied warranties of merchantability or fitness for any particular purpose, and shall in no event be liable for any loss of profit or any other financial or commercial damage, including, but not limited to, special, incidental, consequential, or other damages.

Microsoft Excel is the trademark of Microsoft. All other trademarks are the property of their respective owners.

## Document Purpose / Call for Feedback

The purpose of this document is to assist auditors, fraud examiners, and management in implementing data analysis routines using Microsoft Excel. It is hoped that through the dissemination of this new information that more analysis will be done using audit software to prevent and proactively detect organizational inefficiency, ineffectiveness, and fraud. Please note that although written for auditors, these tests may greatly assist the business community at large as it is currently estimated that there are 400 million users worldwide.

This document is not expected to explain Microsoft Excel concepts at length but rather to provide guidance as to which of the product's features can be used in an audit setting. For more extensive documentation on the use of Microsoft Excel, please see the Help feature in the product and/or the publications section of the [www.microsoft.com](http://www.microsoft.com) website. Another great site, using Excel from a financial perspective, is Business Finance Magazine's area on the use of Excel available at: <http://www.bfmag.com/channels/BrowseArticles.cfm?TopicID=607&CurrChannel=0>

With this document, users can no longer say that audit software is difficult. For now, even a tool as simple as a spreadsheet can and should be used as an audit software. Many readers have noted that this document showed them what is possible with audit software in analogous terms they could understand. With the "curtain exposed" users can now focus their time on the more critical activities of creatively dreaming up new tests and improving their business intelligence.

For more information on the use of audit software, and countless ways of applying it to your business, please see [www.auditsoftware.net/community/](http://www.auditsoftware.net/community/)

If you would like to provide feedback on the document, we welcome and encourage it as we plan to complete later versions. Please provide your feedback via Email at [questions@auditsoftware.net](mailto:questions@auditsoftware.net)

## Limitations of Excel as an Audit Software

Although Microsoft Excel has many powerful features and can take on many of the features of an audit software, it has its own set of limitations that are presented below:

- Can only process 65,536 rows or records of data which may be too small for most organizational databases. Please note however that many report writers (which can handle larger data sets) can define a smaller subset of data for further processing in Excel such as a particular company's division.
- Does not document the auditor's work in easy to access logs for later reference and workpaper storage
- Allows data to be changed in the spreadsheet
- Can only read a small subset of the complete types of data files available in digital format. For example, EBCIDIC files stored in IBM mainframes would need to be converted for use in Excel.
- Has difficulty in performing data management tests such as relating tables. Although it can be accomplished, it is an onerous task.
- Does not have functionality specifically tailored to the auditor. For example, a sample can be calculated in just a couple of clicks with minimal training in a specifically designed audit software. In Excel, it can be done, but it does take some effort and guidance.
- Does not easily apply routines from one data file to other data files whereas audit software can more easily "batch" audit routines for later use on the same file or different data files.

Therefore, for the above features, users should consider more specialized audit software such as ACL Software ([www.acl.com](http://www.acl.com)) or IDEA Software ([www.caseware.com](http://www.caseware.com)). Otherwise, you could consider other database management software such as Microsoft SQL Server ([www.microsoft.com](http://www.microsoft.com)) or Oracle ([www.oracle.com](http://www.oracle.com)). Another option is to review business intelligence tools such as COGNOS ([www.cognos.com](http://www.cognos.com)), Brio ([www.brio.com](http://www.brio.com)) or Business Objects ([www.businessobjects.com](http://www.businessobjects.com)).

# Audit Software Tests

There are two types of audit software tests, those that are analytical in nature and those that are focused on applying analysis to the actual detailed data that are explained below:

## 1. Analytical Tests

*Analytical Tests* - evaluations of financial information made by a study of plausible relationships among both financial and non-financial data to assess whether account balances appear reasonable (AICPA, SAS 56)

This publication explains 11 analytical tests as follows

1. Horizontal Analysis
2. Vertical Analysis
3. Ratios
4. Trend Analysis
5. Performance Measures
6. Statistics
7. Stratifications
8. Aging
9. Benford's Law
10. Regression
11. Monte Carlo Simulation

### 1.1 Horizontal Analysis

Analyzes the increases and decreases in a given balance, normally financial statement items, over two or more periods. This can be completed for the following information using the formula in the most right-hand column (calculating column D):

- Balance sheet
- Income statement
- Budget to actual

|   | A                   | B               | C            | D          | Formula |
|---|---------------------|-----------------|--------------|------------|---------|
| 1 | Asset Description   | This Year       | Last Year    | Difference |         |
| 2 | Cash                | \$1,000         | 900          | (100)      | =B2-C2  |
| 3 | Accounts Receivable | 1,000           | 900          | (100)      | =B3-C3  |
| 4 | Fixed Assets        | 8,000           | 7,200        | (800)      | =B4-C4  |
| 5 | Total Assets        | <b>\$10,000</b> | <b>9,000</b> | (1,000)    | =B5-C5  |

## 1.2 Vertical Analysis

Examines the elements of a financial statement for a single period whereby each balance sheet item is shown as a percentage of the total assets and every income statement item is shown as a percentage of the net sales. This can be completed for the following balance sheet information using the formula in the most right-hand column (calculating column C):

|   | A                   | B             | C            | Formula  |
|---|---------------------|---------------|--------------|----------|
| 1 | Asset Description   | Asset Balance | % of Balance |          |
| 2 | Cash                | \$1,000       | 10%          | =B2/B\$5 |
| 3 | Accounts Receivable | 1,000         | 10%          | =B3/B\$5 |
| 4 | Fixed Assets        | 8,000         | 80%          | =B4/B\$5 |
| 5 | Total Assets        | \$10,000      | 100%         | =B5/B\$5 |

## 1.3 Ratios

One or more balances is compared with one or more other balances such as the relation of total assets to the net sales of an organization. Ratios can be organized into broad categories of “Liquidity/Debt” and “Profitability”.

The Excel calculations for the below ratio tests would be completed using formulas in Excel. For example, for the “Working Capital” (cell B4 below) ratio test, this may be completed as follows:

|   | A                   | B             |
|---|---------------------|---------------|
| 1 | Description         | Asset Balance |
| 2 | Current Assets      | \$1,000       |
| 3 | Current Liabilities | 800           |
| 4 | Working Capital     | 200           |
|   | <b>Formula</b>      | <b>=B2-B3</b> |

A list of common ratio tests are listed below:

Liquidity/Debt – used to measure a company’s ability to pay its vendors or debt obligations in a timely manner.

| Name                          | Description                                                                                               |
|-------------------------------|-----------------------------------------------------------------------------------------------------------|
| Working Capital               | (Current Assets - Current Liabilities)                                                                    |
| Working capital index         | Current year WC - Prior year WC                                                                           |
| Current Ratio                 | (Current Assets / Current Liabilities)                                                                    |
| Days Payable Outstanding      | $(365 / (\text{Sales} / ((\text{Beginning Accounts Payable} + \text{Ending Accounts Payable})/2)))$       |
| Days Sales Outstanding        | $(365 / (\text{Sales} / ((\text{Beginning Accounts Receivable} + \text{Ending Accounts Receivable})/2)))$ |
| Inventory turnover            | $(\text{Cost of Goods Sold} / ((\text{Beginning Inventory} + \text{Ending Inventory})/2))$                |
| Debt to Equity                | Total Debt / Total Stockholders Equity                                                                    |
| Operating Cash Flow           | (Cash Flow From Operations / Current Liabilities)                                                         |
| Cash Flow Interest Coverage   | $((\text{Cash Flow From Operations} + \text{Interest Paid} + \text{Taxes Paid}) / \text{Interest Paid})$  |
| Cash Flow to Capital Expenses | Cash Flow From Operations / Capital Expenses                                                              |

|                          |                                        |
|--------------------------|----------------------------------------|
| Cash Flow to Debt        | Cash Flow From Operations / Total Debt |
| Obsolete Inventory Ratio | Obsolete Inventory / Ending Inventory  |

Profitability – indicate the success of the organization in earning a net return on sales or on an investment.

| Name               | Description                                                                     |
|--------------------|---------------------------------------------------------------------------------|
| Sales growth index | (Current Year Sales / Prior Year Sales)                                         |
| Gross profit       | (Sales – Cost of Goods Sold)                                                    |
| Gross margin       | (Sales – Cost of Goods Sold) / Sales                                            |
| Gross margin index | (Current year Gross Margin / Prior year Gross Margin)                           |
| Stock sales        | (Ending Inventory / Sales)                                                      |
| Return on Equity   | (Net Income / ((Beginning Stockholders Equity + Ending Stockholders Equity)/2)) |

#### 1.4 Trend analysis

Comparing any of the analytical tests (horizontal, vertical, ratio, etc.) described above over two or more periods. Please note that the use of trend analysis is practically a given in doing any audit work as fraud and errors tend to create variances over time which would go undetected if only the single year was being analyzed.

#### 1.5 Performance Measures

The identification of critical success factors that can be tracked over time to assess progress made in achieving specific targets linked to an entity's vision. For example, the below represent a sampling of performance measures that could be used for accounts payable processing:

- Number of invoices processed
- Number of open invoices at period end
- Top 100 vendors purchases
- Average of top 25 max to min payments by vendor ratio
- % of adjustments to invoices processed
- Number of hours overtime worked by staff

The Excel calculations for the above performance measures would be completed using formulas in Excel (i.e., the Sum() function). For more discussion on Excel functions, please see the Extract/Filter section 2.5 of this document.

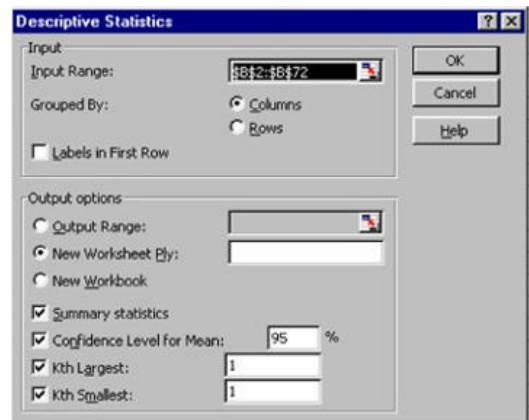
#### 1.6 Statistics

Calculates various statistics. Regarding data such as average, high, low, standard deviation, etc. for a set of numbers. Statistics can be useful in determining the validity of data received and in completing high-level trend analysis.

In order to calculate statistics in Excel, you will first need to add a piece of software that comes with the standard version of Excel. To use this add-in, go to the *Tools* menu in Excel and select *Add-Ins*. Then,

when prompted, select the *Analysis Tool Pack* and press *OK*. This will add to the bottom of your *Tools* menu an item named *Data Analysis*.

Once the information is input, select the *Data Analysis* menu item from the *Tools* menu. Then select the *Analysis Tool Descriptive Statistics* and press *OK* to view the *Descriptive Statistics Dialog Box*:



In the above example, an *Input Range* was entered and the results were pointed to a *New Workbook*. This produces the following results when the *OK* button is selected:

| Column1                 |                    |
|-------------------------|--------------------|
| Mean                    | 208,030.05         |
| Standard Error          | 46,478.93          |
| Median                  | 45,626.97          |
| Mode                    | 100,000.00         |
| Standard Deviation      | 391,638.45         |
| Sample Variance         | 153,380,679,073.67 |
| Kurtosis                | 12.11              |
| Skewness                | 3.26               |
| Range                   | 2,182,906.77       |
| Minimum                 | 1,449.58           |
| Maximum                 | 2,184,356.35       |
| Sum                     | 14,770,133.72      |
| Count                   | 71.00              |
| Largest(1)              | 2,184,356.35       |
| Smallest(1)             | 1,449.58           |
| Confidence Level(95.0%) | 92,699.23          |

## 1.7 Stratifications

Counts the number and dollar value of records of a population falling within specified intervals. Stratifications also provide a useful view into the largest, smallest, and average dollar transactions. An example stratification report by dollar amount is shown below:

| Values               |
|----------------------|
| 100                  |
| 101                  |
| 10                   |
| 1,000                |
| 5,000                |
| <b>6,211 (total)</b> |



| Stratification Report |          |              |
|-----------------------|----------|--------------|
| Strata                | Count    | Dollars      |
| 0 to \$100            | 2        | 110          |
| \$101to \$1,000       | 2        | 1,101        |
| Over \$1,000          | 1        | 5,000        |
| <b>Total</b>          | <b>5</b> | <b>6,211</b> |

To create the above Stratification Report, follow the two-step process below:

Step #1 – Complete a calculation of the strata using a multiple IF function in the right-hand column:

|   | A     | B                                           | C                                                |
|---|-------|---------------------------------------------|--------------------------------------------------|
| 1 | Value | Strata (calculated using formulas at right) | Formula                                          |
| 2 | 100   | 1                                           | =IF(A2>1000,"3",IF(A2>100,"2",IF(A2>0,"1","0"))) |
| 3 | 101   | 2                                           | =IF(A3>1000,"3",IF(A3>100,"2",IF(A3>0,"1","0"))) |
| 4 | 10    | 1                                           | =IF(A4>1000,"3",IF(A4>100,"2",IF(A4>0,"1","0"))) |
| 5 | 1,000 | 2                                           | =IF(A5>1000,"3",IF(A5>100,"2",IF(A5>0,"1","0"))) |
| 6 | 5,000 | 3                                           | =IF(A6>1000,"3",IF(A6>100,"2",IF(A6>0,"1","0"))) |

*Please note that the above IF() function can only handle up to five multiple conditions shown above and therefore, five strata for data stratification purposes.*

Step #2 – Use the strata calculated above to populate the stratification table using the below COUNTIF () and SUMIF () formulas

| Strata | Strata          | Count             | Dollars         |
|--------|-----------------|-------------------|-----------------|
| 1      | 0 to \$100      | =COUNTIF(B2:B6,1) | =SUMIF(B2:B6,1) |
| 2      | \$101to \$1,000 | =COUNTIF(B2:B6,2) | =SUMIF(B2:B6,2) |
| 3      | Over \$1,000    | =COUNTIF(B2:B6,3) | =SUMIF(B2:B6,3) |

### 1.8 Aging

Produces aged summaries of data based on established cutoff dates. This is useful in understanding a process flow over time. An example aging report by dollar amount is shown below:

| Strata       | Count     | Dollars         |
|--------------|-----------|-----------------|
| 0-30 days    | 20        | \$250           |
| 31-60 days   | 10        | \$1500          |
| Over 60 days | 45        | \$30,000        |
| <b>Total</b> | <b>75</b> | <b>\$31,750</b> |

The aging report uses the same concepts explained above for the Stratifications analytical test in section 1.7 of this document.

### 1.9 Benford's Law

Audit technology designed to find abnormal duplications of specific digits, digit combinations, specific numbers, and round numbers in corporate data. Since the objective is to find abnormal duplications, auditors need a benchmark that indicates a normal level of duplication. Benford's Law gives auditors the expected frequencies of the digits in tabulated data. The premise is that we would expect authentic and unmanipulated data to exhibit these patterns. If a data set does not follow these patterns, this may be a cause for auditor concern and to review those. The expected frequencies of Benford's Law for the first digits are:

| Digit | First Digit Frequency |
|-------|-----------------------|
| 0     | -                     |
| 1     | 0.30103               |
| 2     | 0.17609               |
| 3     | 0.12494               |
| 4     | 0.09691               |
| 5     | 0.07918               |
| 6     | 0.06695               |
| 7     | 0.05799               |
| 8     | 0.05115               |
| 9     | 0.04576               |

Step #1 - The first digits of any Excel field can be calculated as follows (using the formula in the right-hand column):

|   | A      | B           |                      |
|---|--------|-------------|----------------------|
| 1 | Values | First Digit | Formula for Column B |
| 2 | 100    | 1           | =LEFT(A2,1)          |
| 3 | 101    | 1           | =LEFT(A3,1)          |
| 4 | 20     | 2           | =LEFT(A4,1)          |
| 5 | 3,000  | 3           | =LEFT(A5,1)          |

|   |       |   |             |
|---|-------|---|-------------|
| 6 | 5,000 | 5 | =LEFT(A6,1) |
|---|-------|---|-------------|

Step #2 – Then use the results of the Left() formulas calculated above to populate the table using the below COUNTIF() function and a simple percentage formula:

|    | D           | E                 | F          | G             |
|----|-------------|-------------------|------------|---------------|
| 1  | First Digit | Count             | Percentage | Benford's Law |
| 2  | 1           | =COUNTIF(B2:B6,1) | =E2/E11    | 0.30103       |
| 3  | 2           | =COUNTIF(B2:B6,2) | =E3/E11    | 0.17609       |
| 4  | 3           | =COUNTIF(B2:B6,3) | =E4/E11    | 0.12494       |
| 5  | 4           | =COUNTIF(B2:B6,4) | =E5/E11    | 0.09691       |
| 6  | 5           | =COUNTIF(B2:B6,5) | =E6/E11    | 0.07918       |
| 7  | 6           | =COUNTIF(B2:B6,6) | =E7/E11    | 0.06695       |
| 8  | 7           | =COUNTIF(B2:B6,7) | =E8/E11    | 0.05799       |
| 9  | 8           | =COUNTIF(B2:B6,8) | =E8/E11    | 0.05115       |
| 10 | 9           | =COUNTIF(B2:B6,9) | =E10/E11   | 0.04576       |
| 11 |             | =Sum(E2:E10)      | =E11/E11   |               |
|    |             |                   |            |               |
|    |             |                   |            |               |

Large deviations between the calculated percentage (Column F) and Benford's Law (Column G) would be investigated. For more information on digital analysis and Benford's Law, please see the following article on ITAudit.org (<http://www.theiia.org/itaudit/index.cfm?fuseaction=forum&fid=95>).

### 1.10 Regression

Regression analysis calculates a dependent variable balance (i.e., net sales) based on various independent variables (i.e., product purchases, inventory levels, number of customers, etc.). Please note that this test generally provides the greatest level of precision because an explicit expectation is formed using all relevant data is incorporated into the model. It also provides a specific precision percentage for each test so that the auditor can assess the reliability of the test.

Given the extensive nature of this topic, this document will not explain regression in depth. Instead, please see the following article which explains step-by-step how to use Excel to complete a regression analysis: <http://www.auditsoftware.net/community/how/tool/tools/regexce.doc>

### **1.11 Monte Carlo simulation**

Monte Carlo allows for the simulation of a balance (i.e., net sales) using estimates where probabilities are given for each estimate. Please note that Monte Carlo simulates the balance estimate thousands of times to arrive at a final estimate with associated precision levels. To understand what Monte Carlo simulation does, think of flipping a coin one hundred times. More than likely, there will be close to 50 heads and 50 tails.

Now, consider a revenue estimate model where there are best case, worst case, and most likely case scenarios given to the perceived market, the number of competitors, the price the market will bear, etc. In contrast with the simple coin flip, a highly advanced probability model can be developed in Monte Carlo tools. In other words, this allows you to flip ten differently weighted coins thousands of times to arrive at a final solution.

Given the extensive nature of this topic, this document will not explain Monte Carlo analysis in depth. For more information on Monte Carlo analysis, please see the following article which explains how add-on software can assist Microsoft Excel in easily calculating a Monte Carlo analysis:

<http://www.auditsoftware.net/community/how/tool/tools/Monte%20Carlo%20Article-Auditors.doc>

## 2. Data Management/Analysis Reports

Data management/analysis reports are those that are run with common audit software yet many can be executed with database management software. These reports are further clarified with specific tests as explained in Chapter 5 of this document. Each type of report is briefly explained below:

| Data Analysis Type             | Description                                                                                                                                                                                                                                                                         |
|--------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Append / Merge              | Combines two files with identical fields into a single file. An example would be to merge two years worth of accounts payable history into one file.                                                                                                                                |
| 2. Calculated Field/ Functions | Created a calculated field (which can use a function such as ABS for the absolute value of the field) using data within the file. For example, the net payroll pay to an employee could be recalculated using the gross pay field and deducting any withholding/taxes.              |
| 3. Cross Tabulate              | Cross Tabulate lets you analyze character fields by setting them in rows and columns. By cross tabulating character fields, you can produce various summaries, explore areas of interest, and accumulate numeric fields.                                                            |
| 4. Duplicates                  | Identifies duplicate items within a specified field in a file. For example, this report could be used to identify duplicate billings of invoices within the sales file.                                                                                                             |
| 5. Extract/Filter              | Extracts specified items from one file and copies them to another file, normally using an “if” or “where” statement. Examples include extracting all balances over a predefined limit.                                                                                              |
| 6. Export                      | Creates a file in another software format (e.g., Excel, Word) for testing. An example would be to export customer address information to Word for “Mail Merge”ing to customer confirmation letters.                                                                                 |
| 7. Gaps                        | Identifies gaps within a specified field in a file. For example, identify any gaps in check sequence.                                                                                                                                                                               |
| 8. Index / Sort                | Sorts a file in ascending or descending order. An example would be sorting a file on social security number to see if any blank or “999999999” numbers exist.                                                                                                                       |
| 9. Join / Relate               | Combines specified fields from two different files into a single file using key fields. This function is used to create relational databases on key fields. For example, the vendor masterfile could be related to the invoice file to obtain address information for each invoice. |
| 10. Sample                     | Creates random or monetary unit samples from a specified population.                                                                                                                                                                                                                |
| 11. Summarize                  | Accumulates numerical values based on a specified key field. An example would be summarizing travel and entertainment expense amounts by employee to identify unusually high payment amounts.                                                                                       |

## 2.1 Append/Merge

Combines two files with identical fields into a single file. An example would be to merge two years worth of accounts payable history into one file as below:

**Current Year File**

|          | <b>A</b>      | <b>B</b>       | <b>C</b>      | <b>D</b>     | <b>E</b>          |
|----------|---------------|----------------|---------------|--------------|-------------------|
| <b>1</b> | <b>Vendor</b> | <b>Invoice</b> | <b>Amount</b> | <b>Check</b> | <b>Check Date</b> |
| <b>2</b> | 11494         | 21597          | 13,466.25     | 6481         | 01/02/1996        |
| <b>3</b> | 1581600       | 29             | 3,562.50      | 6483         | 01/02/1996        |

+

**Prior Year File**

|          | <b>A</b>      | <b>B</b>       | <b>C</b>      | <b>D</b>     | <b>E</b>          |
|----------|---------------|----------------|---------------|--------------|-------------------|
| <b>1</b> | <b>Vendor</b> | <b>Invoice</b> | <b>Amount</b> | <b>Check</b> | <b>Check Date</b> |
| <b>2</b> | 1564189       | 77976          | 48.48         | 6480         | 01/02/1996        |
| <b>3</b> | 25152         | 150842         | 72.56         | 6482         | 01/02/1996        |
| <b>4</b> | 25152         | 150684         | 566.02        | 6482         | 01/02/1996        |



**Combined Current and Prior Year Files**

|          | <b>A</b>      | <b>B</b>       | <b>C</b>      | <b>D</b>     | <b>E</b>          |
|----------|---------------|----------------|---------------|--------------|-------------------|
| <b>1</b> | <b>Vendor</b> | <b>Invoice</b> | <b>Amount</b> | <b>Check</b> | <b>Check Date</b> |
| <b>2</b> | 11494         | 21597          | 13,466.25     | 6481         | 01/02/1996        |
| <b>3</b> | 1581600       | 29             | 3,562.50      | 6483         | 01/02/1996        |
| <b>4</b> | 1564189       | 77976          | 48.48         | 6480         | 01/02/1996        |
| <b>5</b> | 25152         | 150842         | 72.56         | 6482         | 01/02/1996        |
| <b>6</b> | 25152         | 150684         | 566.02        | 6482         | 01/02/1996        |

The Append/Merge in Excel is completed using a simple set of *Copy* and *Paste* commands found in the *Edit* menu. To complete, *Copy* range A1:E2 from the prior year file and *Paste* to a new sheet. Then *Copy* range A2:E3 from the current year file and *Paste* to the same new where the prior data was pasted. This results in the “Combined Current and Prior Year Files” sheet with cells A1 to E6 of current and prior year data.

## 2.2 Calculated Fields/Functions

Created a calculated field using data within the worksheet. For example, the net payroll pay to an employee could be recalculated using the gross pay field and deducting any withholding/taxes as below in column D:

|   | A               | B                | C                      | D              | Formula             |
|---|-----------------|------------------|------------------------|----------------|---------------------|
| 1 | <b>Employee</b> | <b>Gross Pay</b> | <b>Tax &amp; Other</b> | <b>Net Pay</b> | <b>For Column D</b> |
| 2 | Mike            | \$1,000          | 300                    | 700            | =B2-C2              |
| 3 | Jacob           | 1,000            | 280                    | 720            | =B3-C3              |
| 4 | Denise          | 8,000            | 2,700                  | 5,300          | =B4-C4              |
| 5 | <b>Total</b>    | <b>\$10,000</b>  | <b>3,280</b>           | <b>6,720</b>   | <b>=B5-C5</b>       |
|   |                 | =SUM(B1:B4)      | =SUM(C1:C4)            | =SUM(D1:D4)    |                     |

While the =SUM() function was utilized above, there are various other functions available in Excel. Below are a sampling of the most common functions:

**ABS()** - Returns the absolute value of an expression (e.g., ABS(Check\_Amount))

**LOWER()** – Converts a character field to all lower case.

**LEFT()** – Returns the left number of characters (as defined by the user) for the selected field.

**LTRIM()** - Removes all blank spaces from the left of a particular character field, moving all remaining characters to the left.

**NPV()** - Calculates the net present value of an investment by using a discount rate and a series of future payments (negative values) and income (positive values).

**ROUND()** – rounds a cell's value to a specified number of digits.

**SUM()** – Adds the numbers in a range of cells.

**UPPER()** - Converts a character field to all upper case.

**VALUE ()** - Converts a character expression or field into the numeric equivalent.

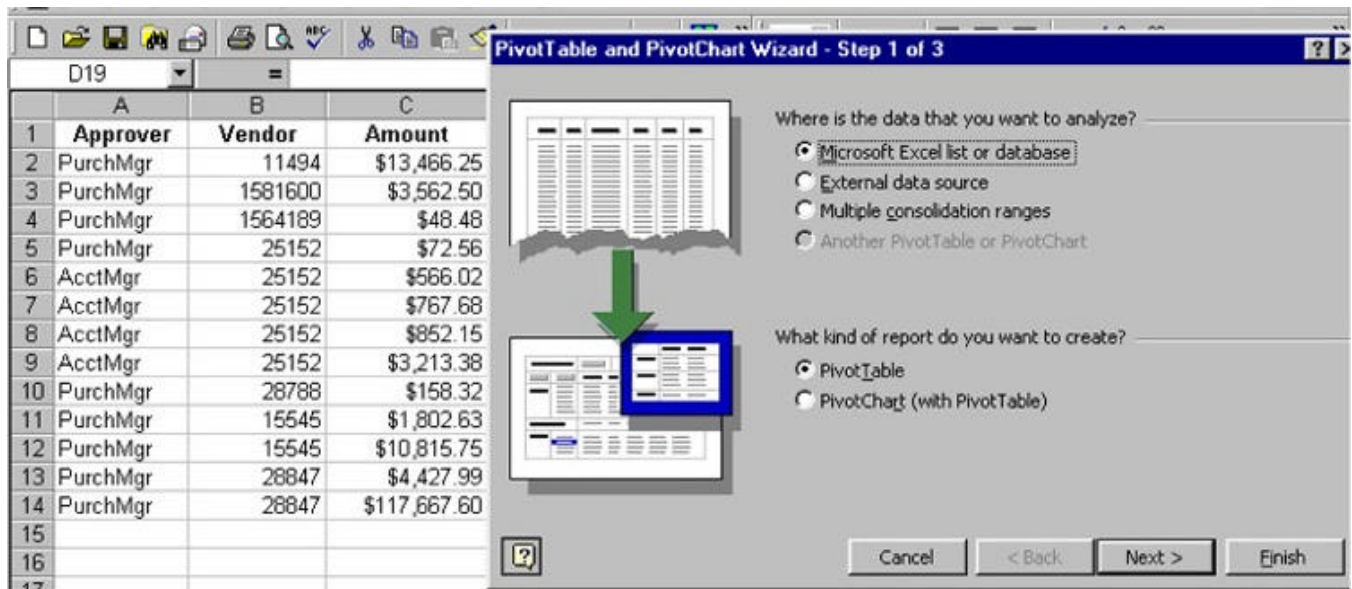
**WEEKDAY()** - Returns the day of the week corresponding to a date. The day is given as an integer, ranging from 1 (Sunday) to 7 (Saturday), by default.

**WEEKNUM()** - Returns a number that indicates where the week falls numerically within a year.

## 2.3 Cross Tabulate

Cross Tabulate lets you analyze character fields by setting them in rows and columns. By cross tabulating character fields, you can produce various summaries, explore areas of interest, and accumulate numeric fields.

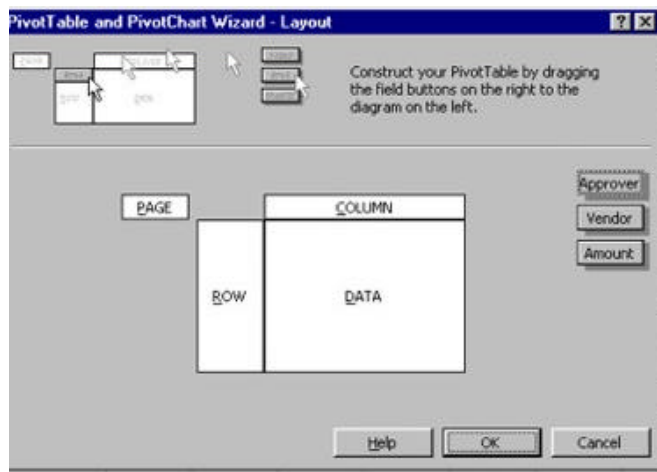
Excel effectuates cross-tabulation through its *Pivot Tables* feature explained below for the analysis of vendor payments in relation to approvers. To open the dialog box, select *Data* from the menu bar and *PivotTable and Pivot Chart Report...*:



The PivotTable Wizard will walk you through the selection of the data range and will result in the following dialog box:



At this point, select the *Layout* button in order to establish the columns and rows of the pivot table. By starting with picture #1 below, the fields on the right (Approver, Vendor, and Amount) can be selected and dragged as shown in picture #2 below:



#1



#2

Please note that in the above box that the Sum of Amount field may start out as the Count of Amount when you first attempt this maneuver. To change it from Count to Sum (so that the total dollars can be calculated) drag it over and then double click on the Count of Amount. From here, simply change it to Sum. To close the Layout window, select OK and select Finish from the previous window.

Once complete, select the *OK* and then the *Finish* buttons to arrive at the below cross-tabulation:

| Sum of Amount | Approver |           |             |
|---------------|----------|-----------|-------------|
| Vendor        | AcctMgr  | PurchMgr  | Grand Total |
| 11494         |          | 13466.25  | 13466.25    |
| 15545         |          | 12618.38  | 12618.38    |
| 25152         | 5399.23  | 72.56     | 5471.79     |
| 28788         |          | 158.32    | 158.32      |
| 28847         |          | 122095.59 | 122095.59   |
| 1564189       |          | 48.48     | 48.48       |
| 1581600       |          | 3562.5    | 3562.5      |
| Grand Total   | 5399.23  | 152022.08 | 157421.31   |

Please note that once this summary report is created, it can be filtered by the dropdown arrows (see above picture and section 2.6 of this document for more on Excel filters) so that purchases with a specific vendor could be summarized and reviewed in relation to the approvers of those purchases.

## 2.4 Duplicates

Identifies duplicate items within a specified field in a file. For example, the below worksheet could be established to identify duplicate billings of invoices within the sales file where the same customer, invoice, and amount are the same:

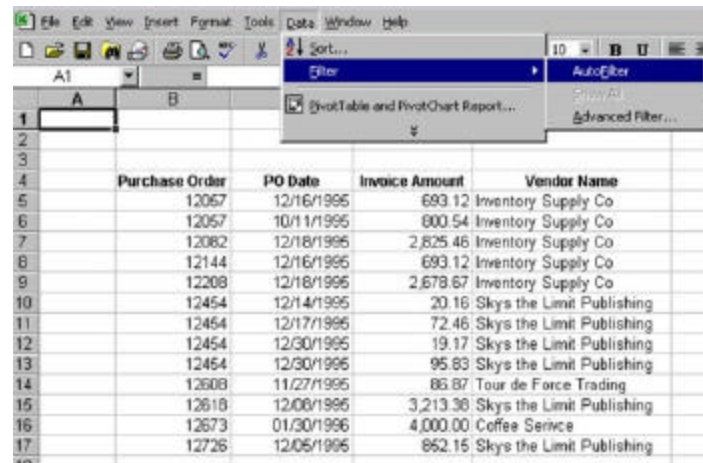
|   | A       | B               | C         | D                |                                     |
|---|---------|-----------------|-----------|------------------|-------------------------------------|
| 1 | Invoice | Customer        | Amount    | Duplicate<br>Y/N | Formula for Column D                |
| 2 | 77524   | Sky's the Limit | \$800.00  |                  |                                     |
| 3 | 77525   | Sky's the Limit | \$1752.23 | N                | =IF(AND(A3=A2,B3=B2,C3=C2),"Y","N") |
| 4 | 77525   | Sky's the Limit | \$1752.23 | Y                | =IF(AND(A4=A3,B4=B3,C4=C3),"Y","N") |
| 5 | 77528   | Sky's the Limit | 922.00    | N                | =IF(AND(A5=A4,B5=B4,C5=C4),"Y","N") |
| 6 | 77535   | Sky's the Limit | 750.00    | N                | =IF(AND(A6=A5,B6=B5,C6=C5),"Y","N") |

Essentially, the *IF()* and *AND()* functions are used to create a column D which calculates whether the item is a duplicate.

***Please note that the AND() function above can accept up to 30 expressions (three are used in the above AND() function).***

## 2.5 Extract/Filter

Extracts specified items from one file and copies them to another file, normally using an “if” or “where” statement. Examples include extracting all balances over a predefined limit. This is most easily accomplished using the *Auto Filter* command under the *Filter* option of the *Data* menu.



For example, let's presume you needed to identify all payments to “Sky's the Limit Publishing” over \$1,000.

**Step #1** – Highlight the row with each column's description (as the in below picture) and select the *Auto Filter* command under the *Filter* option of the *Data* menu (as the in above picture) to be presented with down arrows next to each column's description (as the in below picture):

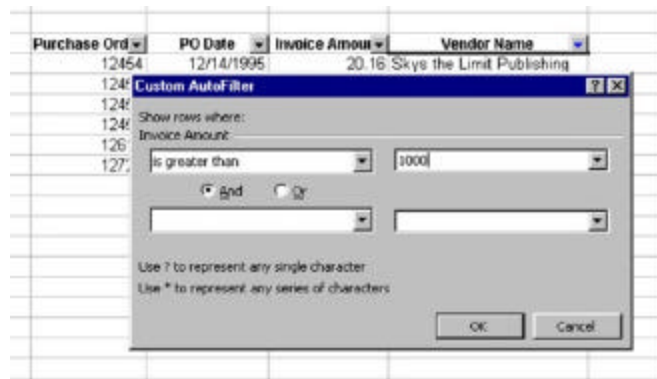
|    | A | B            | C          | D              | E                         | F |
|----|---|--------------|------------|----------------|---------------------------|---|
| 1  |   |              |            |                |                           |   |
| 2  |   |              |            |                |                           |   |
| 3  |   |              |            |                |                           |   |
| 4  |   | Purchase Ord | PO Date    | Invoice Amount | Vendor Name               |   |
| 5  |   | 12057        | 12/16/1995 | 693.12         | Inventory Supply Co       |   |
| 6  |   | 12057        | 10/11/1995 | 800.54         | Inventory Supply Co       |   |
| 7  |   | 12082        | 12/18/1995 | 2,825.46       | Inventory Supply Co       |   |
| 8  |   | 12144        | 12/16/1995 | 693.12         | Inventory Supply Co       |   |
| 9  |   | 12208        | 12/18/1995 | 2,678.67       | Inventory Supply Co       |   |
| 10 |   | 12454        | 12/14/1995 | 20.16          | Skys the Limit Publishing |   |
| 11 |   | 12454        | 12/17/1995 | 72.46          | Skys the Limit Publishing |   |

**Please note that in the above step that it is important there NOT be any blank lines between the column heading and the underlying data. Otherwise, the Filter will not operate effectively.**

**Step #2** – By selecting the down arrow next to the field **Vendor Name**, you will be presented with a list of all vendor names within that particular column. Using this, the “Sky's the Limit” vendor can be selected:

| Purchase Ord | PO Date    | Invoice Amount | Vendor Name               |
|--------------|------------|----------------|---------------------------|
| 12454        | 12/14/1995 | 20.16          | (All)                     |
| 12454        | 12/17/1995 | 72.46          | (Top 10...)               |
| 12454        | 12/30/1995 | 19.17          | (Custom...)               |
| 12454        | 12/30/1995 | 95.63          | Coffee Service            |
| 12618        | 12/08/1995 | 3,213.38       | Inventory Supply Co       |
| 12726        | 12/05/1995 | 852.15         | Skys the Limit Publishing |
|              |            |                | Tour de Force Trading     |
|              |            |                | Skys the Limit Publishing |

This will filter all items in the worksheet with “Sky’s the Limit” in the vendor name field. Then, the down arrow next to the **Invoice Amount** column can be selected and this time, the (*Custom...*) item can be selected (see above picture) to be presented with the following dialog box:



Working with the above box, the *is greater than* option can be selected and an amount 1000 can be entered to filter the invoice amount field for anything greater than \$1,000 where “Sky’s the Limit” is the vendor, resulting in the below:

| Purchase Ord | PO Date    | Invoice Amount | Vendor Name               |
|--------------|------------|----------------|---------------------------|
| 12618        | 12/08/1995 | 3,213.36       | Skys the Limit Publishing |

Please Note the following when using the Auto Filter feature in Excel:

- Using the (*Custom...*) feature, you can select up to three expressions with *AND* or *OR* statements in between each expression.
- Do not have any rows between the description row and the data or the *Auto Filter* feature will not operate properly.
- Do not have any blank rows within the worksheet data or the *Auto Filter* feature will not operate properly.
- To remove the Auto Filter feature, select *Auto Filter* command under the *Filter* option of the *Data* menu which will deselect the command.

## 2.6 Export

Creates a file in another software format for testing. An example would be to export customer address information to a tab delimited file so that it may be imported within another audit software. All exporting from Excel is effectuated through the *Save As* feature in the *File* menu. Common options for exporting data include:

- Comma delimited
- Dbase
- Tab delimited
- Web page

## 2.7 Gaps

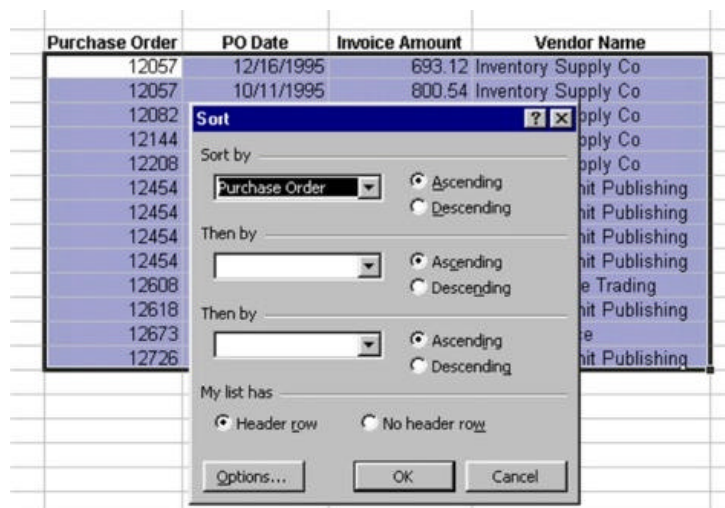
Identifies gaps within a specified field in a file. For example, to identify any gaps in check sequence in the below table, the following formula (see cell C3) would be entered and copied down for the entire length of data in the worksheet (in this spreadsheet this would start at B3 and copied down to B6):

|   | A            | B    | C                         |
|---|--------------|------|---------------------------|
| 1 | Check Number | Gaps | Formula for Column B      |
| 2 | 77524        | 0    |                           |
| 3 | 77525        | 0    | =IF(A3-A2<>1,(A3-A2)-1,0) |
| 4 | 77527        | 1    | =IF(A4-A3<>1,(A4-A3)-1,0) |
| 5 | 77528        | 0    | =IF(A5-A4<>1,(A5-A4)-1,0) |
| 6 | 77535        | 6    | =IF(A6-A5<>1,(A6-A5)-1,0) |

**\*\*\*Please note that before you run the GAPS test above, you need to sort your file on the field being tested for gaps in Ascending order (see 2.8 Index/Sort Below)**

## 2.8 Index / Sort

Sorts a file in ascending or descending order. An example would be sorting a file on social security number to see if any blank or “999999999” numbers exist. To sort a file in Excel, the range of data must first be highlighted. Then, the *Sort* command is selected from the *Data* menu item to be presented the following self-explanatory dialog box:



## 2.9 Join / Relate

Combines specified fields from two different files into a single file using key fields. This function is used to create relational databases on key fields and mainly uses the *VLOOKUP()* function in Excel. Please note that of all of the Excel uses as an audit tool, the joining of files is most difficult and better completed by an audit or database software. However, simple joining of files is possible and explained further below.

For example, to join the vendor masterfile to the employee masterfile on the address column, the following functions (column E and G) need to be defined (which are further explained below the picture):

|    | A                     | B | C                       | D                     | E                        | F             | G                  | H |
|----|-----------------------|---|-------------------------|-----------------------|--------------------------|---------------|--------------------|---|
| 1  |                       |   |                         |                       |                          |               |                    |   |
| 2  |                       |   |                         | <b>Most Likely</b>    |                          |               |                    |   |
| 3  | <b>Vendor Address</b> |   | <b>Employee Address</b> | <b>Vendor Address</b> | <b>Formula</b>           | <b>Match?</b> | <b>Formula</b>     |   |
| 4  | 120 Berger Place      |   | 20816 HardwoodCourt     | 145 Ridge Road        | =VLOOKUP(C4,A\$3:A\$8,1) | N             | =IF(D4=C4,"Y","N") |   |
| 5  | 145 Ridge Road        |   | 22 Waverly Place        | 22 Waverly Place      | =VLOOKUP(C5,A\$3:A\$8,1) | Y             | =IF(D5=C5,"Y","N") |   |
| 6  | 22 Waverly Place      |   | 35 Waldorf Avenue       | 2412 Maple Ave        | =VLOOKUP(C6,A\$3:A\$8,1) | N             | =IF(D6=D6,"Y","N") |   |
| 7  | 2300 Pilgrim Lane     |   | 4 Meadowland Court      | 2412 Maple Ave        | =VLOOKUP(C7,A\$3:A\$8,1) | N             | =IF(D7=D7,"Y","N") |   |
| 8  | 2412 Maple Ave        |   | 42 Park Avenue          | 2412 Maple Ave        | =VLOOKUP(C8,A\$3:A\$8,1) | N             | =IF(D8=D8,"Y","N") |   |
| 9  |                       |   |                         |                       |                          |               |                    |   |
| 10 |                       |   |                         |                       |                          |               |                    |   |
| 11 |                       |   |                         |                       |                          |               |                    |   |

Step 1: Join the vendor and employee address columns using the formulas listed in column D. The *VLOOKUP()* function takes an address in column C (employee address) and matches it to Column A (vendor address). Please note that Excel looks for the most precise value and will not leave the calculated column blank. Therefore,

|   | A                     | B | C                       | D                                 |                             |
|---|-----------------------|---|-------------------------|-----------------------------------|-----------------------------|
| 1 |                       |   |                         |                                   |                             |
| 2 |                       |   |                         |                                   |                             |
| 3 | <b>Vendor Address</b> |   | <b>Employee Address</b> | <b>Most Likely Vendor Address</b> | <b>Formula for Column D</b> |
| 4 | 120 Berger Place      |   | 20816 HardwoodCourt     | 145 Ridge Road                    | =VLOOKUP(C4,A\$3:A\$8,1)    |
| 5 | 145 Ridge Road        |   | 22 Waverly Place        | 22 Waverly Place                  | =VLOOKUP(C5,A\$3:A\$8,1)    |
| 6 | 22 Waverly Place      |   | 35 Waldorf Avenue       | 2412 Maple Ave                    | =VLOOKUP(C6,A\$3:A\$8,1)    |
| 7 | 2300 Pilgrim Lane     |   | 4 Meadowland Court      | 2412 Maple Ave                    | =VLOOKUP(C7,A\$3:A\$8,1)    |
| 8 | 2412 Maple Ave        |   | 42 Park Avenue          | 2412 Maple Ave                    | =VLOOKUP(C8,A\$3:A\$8,1)    |
|   |                       |   |                         |                                   |                             |
|   |                       |   |                         |                                   |                             |

**Step 2:** Now that the most likely vendor address matching to the employee address has been calculated, a simple IF() function is written to identify any true matches between the vendor and employee address files.

| Match? | Formula            |  |
|--------|--------------------|--|
| N      | =IF(D4=C4,"Y","N") |  |
| Y      | =IF(D5=C5,"Y","N") |  |
| N      | =IF(D6=D6,"Y","N") |  |
| N      | =IF(D7=D7,"Y","N") |  |
| N      | =IF(D8=D8,"Y","N") |  |



The end result in the above analysis is that cell C3 (Waverly Place) has a match between the vendor and employee address files which is noted with a “Y” in the above picture.

**Key Note on Joining Files Using VLOOPKUP:** Please note that matching on an entire address field, as in the above example, may not be practical given address fields in the vendor and employee masterfile may be entered differently. For example “123 Maple Street” in the vendor file may be “123 Maple St.” in the employee file and therefore not show up as a match.

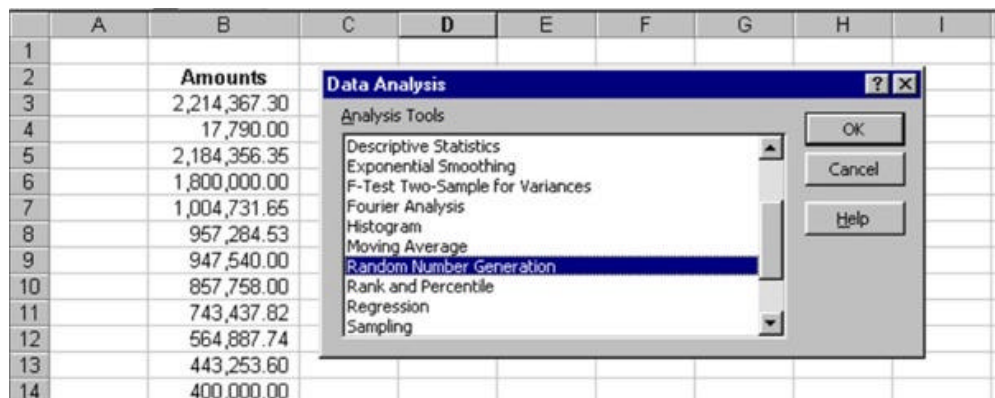
To avoid the above problem, the LEFT() function may be used to select say the first 8 characters of the address fields. For example, LEFT(A4,8) in the above example would return “123 Berg”. This calculated field could be made for both the vendor and employee address columns and a VLOOPKUP used on the newly calculated column to get a match on the first eight characters of the address field.

## 2.10 Sample

Creates random or monetary unit samples from a specified population. In Excel, a random sample can be easily completed as follows for a sampling of data below:

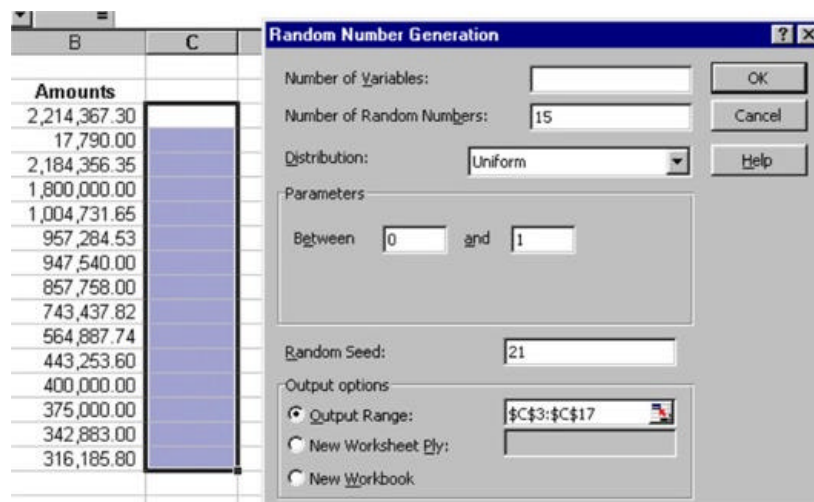
Step #1 - In order to calculate statistics in Excel, you will first need to add a piece of software that comes with the standard version of Excel. To use this add-in, go to the *Tools* menu in Excel and select *Add-Ins*. Then, when prompted, select the *Analysis Tool Pack* and press *OK*. This will add to the bottom of your *Tools* menu an item named *Data Analysis*.

Step #2 – Select the *Random Number Generator* option as seen below:



Step #3 – Complete the resulting dialog box noting the following:

- The *Number of Random Numbers* should equate to the number of rows in the file. In my example (below picture) there are fifteen items in the row.
- The *Distribution* should be set to *Uniform*.
- A *Random Seed* should be entered which will affect the generation of random numbers. It is suggested that a different seed number be entered for each sample selected.
- An *Output Range* must be designated which is suggested to be next to the data to be sampled as done in the below dialog box with the results seen on the following page.



Step #4 – Once the random numbers have been generated and placed next to the data for sampling, the worksheet of data would need to be sorted using the *Sort* command in Excel (see section 2.8 of this document).

| B              | C        |
|----------------|----------|
| <b>Amounts</b> |          |
| 2,214,367.30   | 0.003265 |
| 17,790.00      | 0.123905 |
| 2,184,356.35   | 0.096713 |
| 1,800,000.00   | 0.495804 |
| 1,004,731.65   | 0.880551 |
| 957,284.53     | 0.876553 |
| 947,540.00     | 0.249977 |
| 857,758.00     | 0.109256 |
| 743,437.82     | 0.900693 |
| 564,887.74     | 0.99414  |
| 443,253.60     | 0.083834 |
| 400,000.00     | 0.8858   |
| 375,000.00     | 0.222999 |
| 342,883.00     | 0.612629 |
| 316,185.80     | 0.849055 |

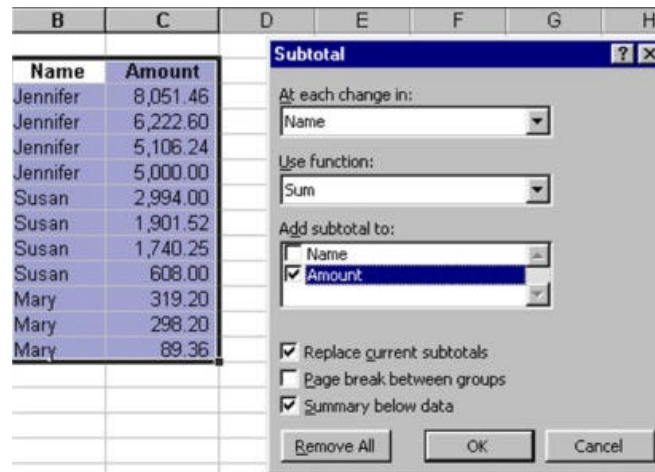
Final Result – With the final results sorted in ascending order (see below picture) a sample can be selected starting with the first item and working downwards. For example, if five items were to be sampled, the first five items would represent the random sample (see below picture):

| B              | C             | D                 |
|----------------|---------------|-------------------|
| <b>Amounts</b> | <b>Random</b> | <b>Selections</b> |
| 2,214,367.30   | 0.003265      | 1                 |
| 443,253.60     | 0.083834      | 2                 |
| 2,184,356.35   | 0.096713      | 3                 |
| 857,758.00     | 0.109256      | 4                 |
| 17,790.00      | 0.123905      | 5                 |
| 375,000.00     | 0.222999      |                   |
| 947,540.00     | 0.249977      |                   |
| 1,800,000.00   | 0.495804      |                   |
| 342,883.00     | 0.612629      |                   |
| 316,185.80     | 0.849055      |                   |
| 957,284.53     | 0.876553      |                   |
| 1,004,731.65   | 0.880551      |                   |
| 400,000.00     | 0.8858        |                   |
| 743,437.82     | 0.900693      |                   |
| 564,887.74     | 0.99414       |                   |

## 2.11 Summarize

Accumulates numerical values based on a specified key field.

For example one potential test could be summarizing travel and entertainment expense amounts by employee to identify unusually high payment amounts. To complete this test, first sort the data using the Sort command explained above. Then, highlight the data to subtotal and then select *Subtotals* from the *Data* menu item to be presented with the following dialog box:



Using this dialog box, the columns to subtotal can be selected and, when finished, the OK button can be selected to complete the summarizations below. Please note that Options 1, 2, and 3 relate to the button on the left-top side of each individual picture below to be presented with a grand total (Option 1), a total by employee (Option 2) or a detail listing with subtotals (Option 3):

### Option 1

|    | A | B           | C         |
|----|---|-------------|-----------|
| 1  |   |             |           |
| 2  |   | Name        | Amount    |
| 17 |   | Grand Total | 32,330.83 |

### Option 2

|    | A | B              | C         |
|----|---|----------------|-----------|
| 1  |   |                |           |
| 2  |   | Name           | Amount    |
| 7  |   | Jennifer Total | 24,380.30 |
| 12 |   | Susan Total    | 7,243.77  |
| 16 |   | Mary Total     | 706.76    |
| 17 |   | Grand Total    | 32,330.83 |

### Option 3

|    | A | B              | C         |
|----|---|----------------|-----------|
| 1  |   |                |           |
| 2  |   | Name           | Amount    |
| 3  |   | Jennifer       | 8,051.46  |
| 4  |   | Jennifer       | 6,222.60  |
| 5  |   | Jennifer       | 5,106.24  |
| 6  |   | Jennifer       | 5,000.00  |
| 7  |   | Jennifer Total | 24,380.30 |
| 8  |   | Susan          | 2,994.00  |
| 9  |   | Susan          | 1,901.52  |
| 10 |   | Susan          | 1,740.25  |
| 11 |   | Susan          | 608.00    |
| 12 |   | Susan Total    | 7,243.77  |
| 13 |   | Mary           | 319.20    |
| 14 |   | Mary           | 298.20    |
| 15 |   | Mary           | 89.36     |
| 16 |   | Mary Total     | 706.76    |
| 17 |   | Grand Total    | 32,330.83 |

## About the Author

Rich Lanza (CPA, PMP) is a Manager of Internal Audit at a Fortune 200 retailer, where he focuses mainly on using computer assisted audit tools to improve business intelligence, increase efficiencies, and identify multi-million dollar bottom-line savings.

Rich is a leading authority on the use of data extraction/analysis technology. He has devoted himself to freely providing information on the topic as his mission is to help auditors get recognized for their bottom-line results. To that end, he founded *AuditSoftware.Net* ([www.auditsoftware.net/community](http://www.auditsoftware.net/community)) that works to increase organizational benefits from the use of audit software. The free site provides tools, case studies, a vendor discount program, newsletter, and a discussion area to visitors. The site recently announced project communities whereby anyone interested can share and collaborate on practical audit software tools and solutions. Rich has also written numerous articles, software products, and four test set books on how to practically apply ACL and IDEA software (<http://www.ekaros.ca/publications.html>).

Rich headed the Program Management Office, reporting to the Chief Operating Officer at the American Institute of Certified Public Accountants, where he sewed a culture of project management into the fabric of the organization. On his last major project, Rich worked to coordinate various fraud-reduction initiatives to maximize their benefit for the AICPA and financial markets. Rich has drafted many of the project standards being used at the AICPA to manage projects and has implemented a project portfolio management process. He has also established a six-month intense training program for project managers. Prior to his program management work, Rich was a program/project manager in many Web and technology projects including leading the Y2K project.

Prior to joining the AICPA, Rich served as a Vice President-Audit Technology at AuditWatch where he was an ACL® and IDEA™ trainer. He has also been an Internal Audit Manager for Lafarge and Disclosure and has served in the audit department at KPMG Peat Marwick. Rich received his undergraduate degree in public accounting from Pace University. He is a past President of the Northern Virginia Chapter of The Institute of Internal Auditors, and is a certified member of both the AICPA and PMI. He currently also maintains the New York/New Jersey Program Management Office Special Interest Group for PMI.

In 2003, for his data analysis work in the area of fraud detection, he was awarded the Outstanding Achievement in Commerce Award from the Association of Certified Fraud Examiners and in he won the Second Annual ACL Users Challenge. His interests include magic, and collecting gadgets.

The author's opinions expressed in this publication are his own and do not necessarily represent the policies or positions of his employer.

Rich can be reached through the following means:

E-mail: [questions@richlanza.com](mailto:questions@richlanza.com)

Phone: +1-973-828-0239

Websites: [www.richlanza.com](http://www.richlanza.com)  
[www.auditsoftware.net/community](http://www.auditsoftware.net/community)