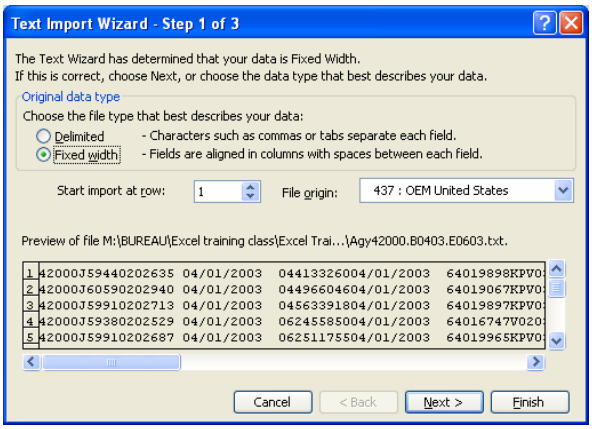
Excel Tips and Tricks to Find Improper Payments

To obtain a copy of the supporting documents referred to in these instructions, send an email to the Outreach Training Program - Outreach@osc.state.ny.us

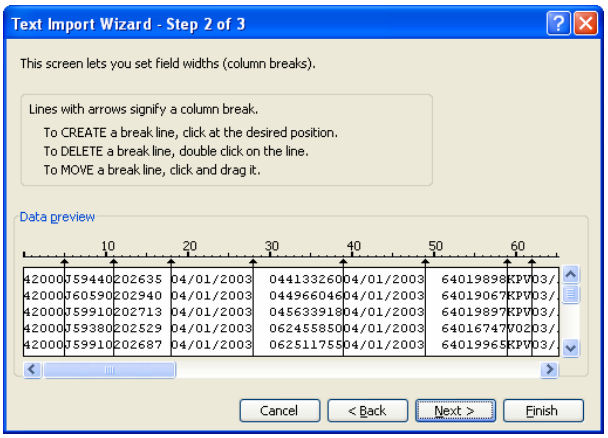
IMPORT DATA FROM TEXT FILE

Open Excel
File → Open
Bottom of "Open" screen change "Files of type" to All Files
Find the folder containing the Text file
Double click on the file to open

Depending on text file type, select either Delimited or Fixed Width. In this case, the file is Fixed Width.



Single-click to add markers or click-and-hold to drag markers to the desired position. Double click to remove markers. The lines should be at position numbers:
5,11,18,28,39,49,59,62,72,75,89,99,116,126,142,149,169,174,184



Define the field types and click Finish
Save the file as an Excel workbook

VERIFY THE DATA

Verify Payment Amount:
Use the **SUM** function
Should have $52,812,312.74

Verify number of records:
Use the **COUNT** function
Should have 4,448

Verify Dates:
Highlight all records
Data → Sort → Payment Date
4/1/2003 – 6/30/2003

UNDERSTANDING YOUR DATA

FILTERS

Quick Filter:

Highlight column names
Data → Filter → Auto Filter

Sum and count for filtered data:

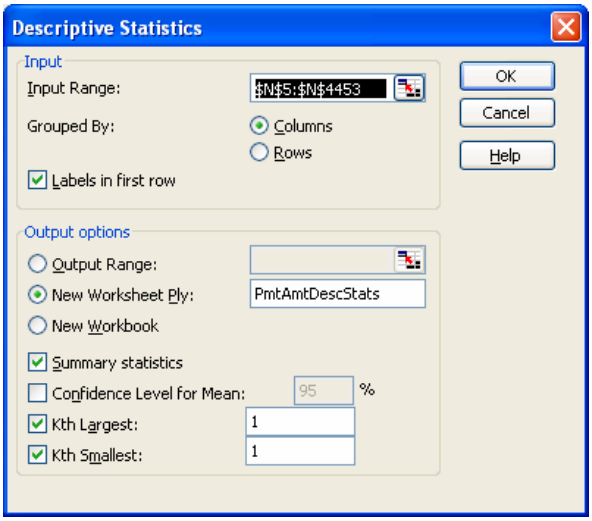
Use the **SUBTOTAL** function
Sum (9) – Use the $ sign to lock in the cell references, in order to copy the formula down to the next cell.
Count (3) - Change the Function number to 3

Now you can apply filters and the **SUBTOTAL** function will total and count only the non-filtered records.

Descriptive Statistics

Microsoft Excel provides a set of data analysis tools, called the Analysis ToolPak that you can use to save steps when you develop complex statistical analyses.
Click Tools→ Add_Ins→Check Analysis ToolPak
Click Tools→Data Analysis→Descriptive Statistics

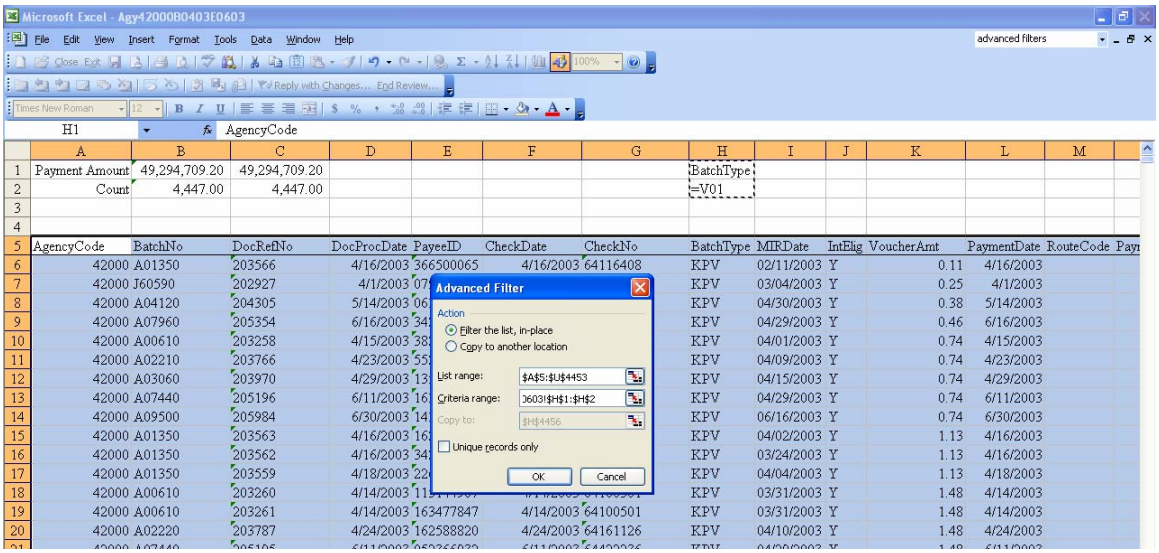
Enter the range for a numeric field (PaymentAmt)
Check the box "Labels in First Row"
Select New Worksheet Ply (name optional)
Check the Summary Statistics checkbox
Check both Kth Largest and Kth Smallest, enter "1" in the field



SUMMARIZING DATA

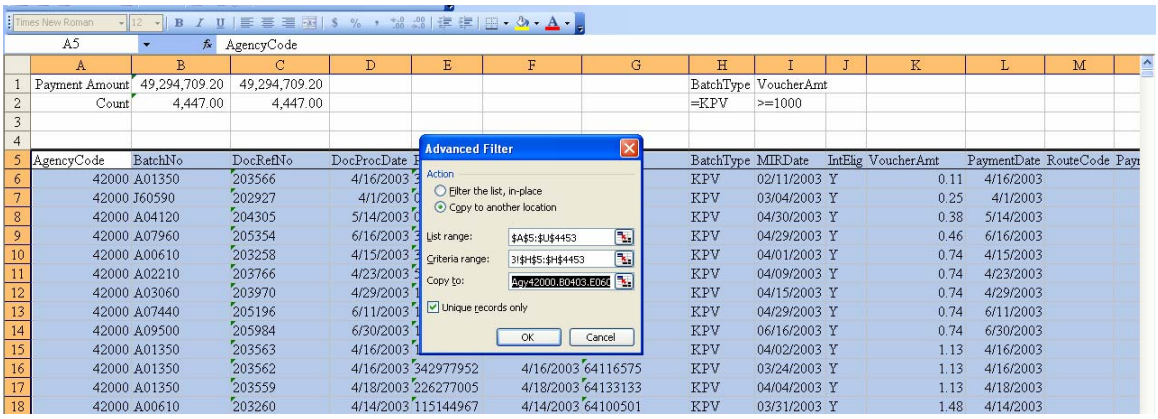
ADVANCED FILTERS

Filter your data on multiple fields and multiple criteria
In the cell H1, enter the field name exactly how it appears in the data (BatchType)
In the cell H2, enter the criteria as =”=KPV”
Data→Filter→Advanced Filter
Select the radio button Filter the List, in-place
The list range is your entire table ($A$5:$U$44453)
The Criteria range is the cells you created ($H$1:$H$2)



Multiple Criteria
In I1 enter the field name “VoucherAmt”
In I2, enter the amount =”>=1000”
Data→Filter→Advanced Filter
Select the radio button Filter the List, in-place
The list range is your entire table ($A$5:$U$44453)
The Criteria range is the cells you created ($H$1:$I$2)

Filtering for unique records
Data→Filter→Advanced Filter
Select the radio button “Copy to another location”
Select the list range for BatchTypes (H6:H4453)
Copy to a cell at the bottom of the sheet

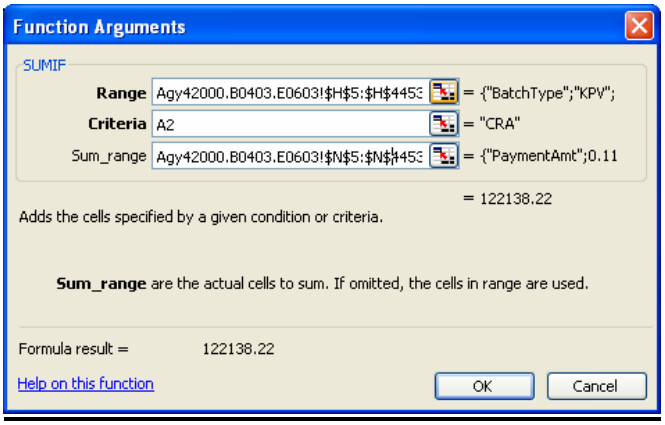


Create a new sheet and name it BatchTypes
Copy the filtered Batch Types to the “BatchTypes” worksheet
Sort Ascending and delete any duplicates

CALCULATED FIELDS

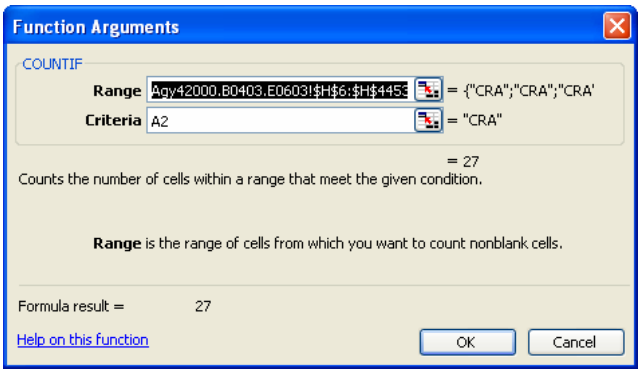
SumIF
Calculate PaymentAmt for each BatchType
Click the Insert Function button
Range – Where to look.
Criteria – What do you want to look for.
Sum_Range – What to add.

Add $ signs before the Column letter and Row number to lock in the range



CountIF

Calculate number of payments for each BatchType
Click the Insert Function button



CONDITIONAL FORMULAS

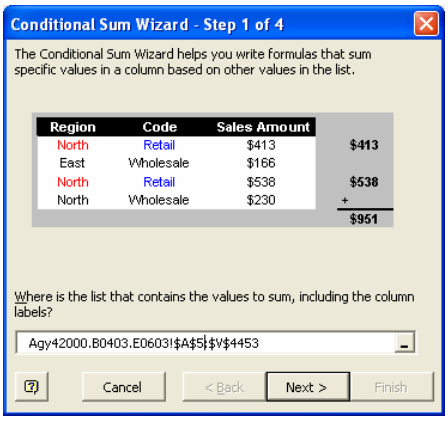
Summarize data based on multiple criteria

In BatchType add columns for April, May and June
In the Agy42000.B0403.E0603 worksheet add a column for PaymentMonth
Use the **Month** function to return the month number

Conditional Sum

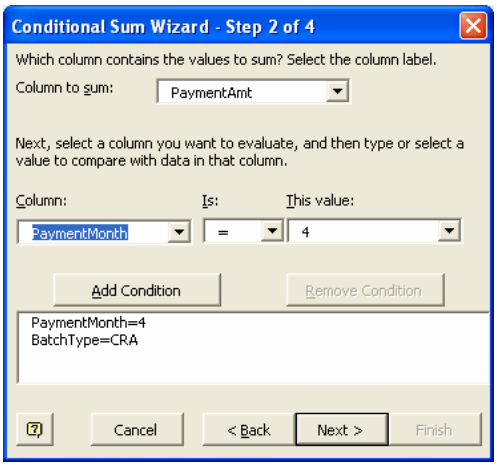
Tools → Add-ins→Conditional Sum Wizard
Tools → Conditional Sum

Select the entire Payment History table

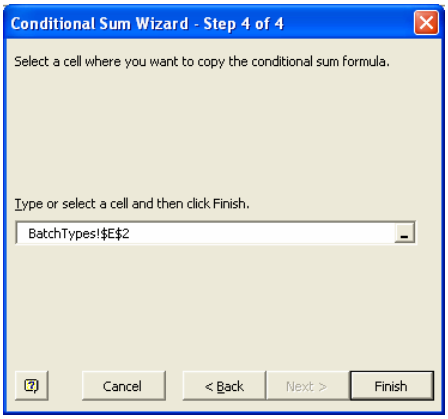


Column to Sum – PaymentAmount
Add the conditions:

BatchType = CRA
PaymentMonth = 4



Select the cell to put the result



Press F2 to change the formula so it looks in cell A2 for the BatchType. Add $ sign in front of the letter A.

To lock in the formula, press Ctrl, Shift and Enter
Copy and paste the formula from E2 to F2.
Change the month to 5 and hit Ctrl, Shift and Enter to lock it in
Copy and paste the formula down
Repeat for June

PIVOT TABLE

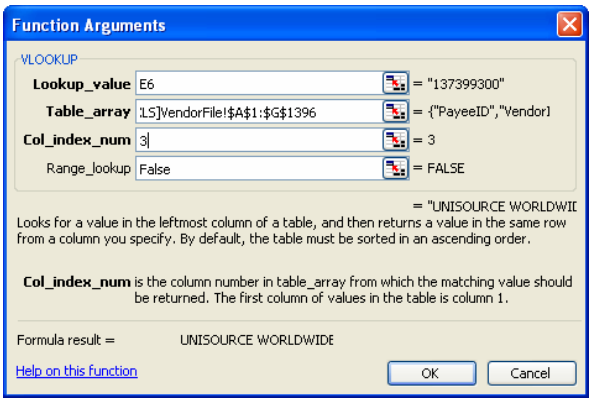
PivotTable reports are interactive tables that help you to organize and compare data so that you can see patterns, trends, and relationships.

Create new worksheet called "Pivot"
Data → Pivot Table and Pivot Chart Report
Select the "Microsoft Office Excel list or database" radio button
Select the PivotTable radio button
Click Next
Range – is the entire table from Agy42000.B0403.E0603 including column headings
Click the Layout button
Drag the PayeeID into the Row field
Drag the PaymentMonth into the Column field
Drag the PaymentAmt into the Data field twice
Double click on the second PaymentAmt to select Count (if necessary)
Click OK then Finish

VLOOKUP

Searches for a value in the first column of a table array and returns a value in the same row from another column in the table array
You will need to identify the key fields (in this case it's PayeeID)

In the Agy42000.B0403.E0603 table, insert a column to the right of the key field (PayeeID) and name it PayeeName
Open the DOESVendor File
The key field (PayeeID) needs to be in the ColumnA
In the Agy42000.B0403.E0603, click on the first blank record in the PayeeName field.
Click on the Insert Function button and search for **VLOOKUP**
Select **VLOOKUP** and click OK
Lookup_value – is the first PayeeID in the Agy42000.B0403.E0603 table
Table_array – is the entire DOESVendor File table (make sure you have the $ signs to lock in the values)
Col_index_Num – is the column **number** (3)of the matching value
Range_lookup – Exact Match = **False**. Closest Match = **True**



Click OK
Copy the formula down

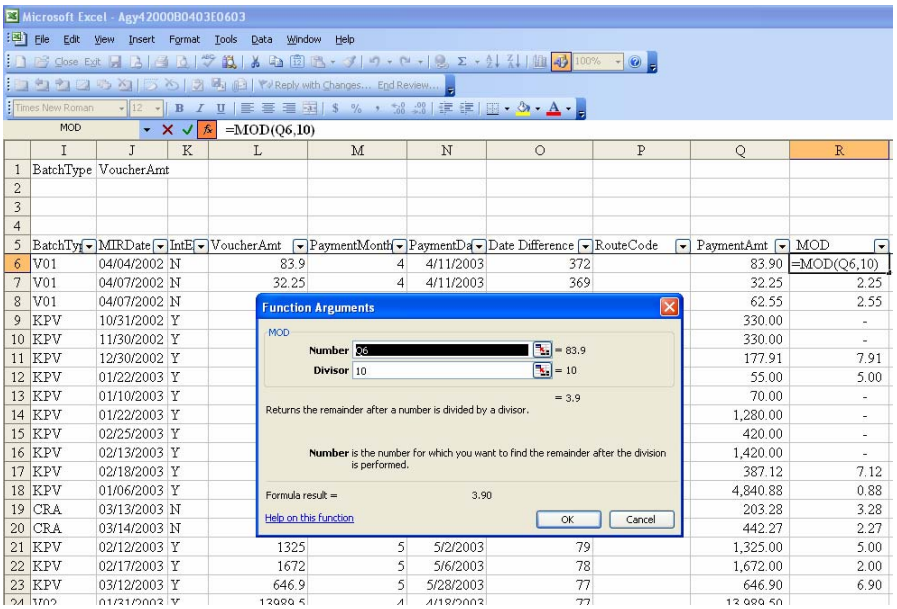
Go to the Pivot Table
Click Refresh
Add PayeeName to the Row field

Copy and paste the PayeeName to lock in the values

USEFUL FUNCTIONS / IMPROPER PAYMENT TESTS

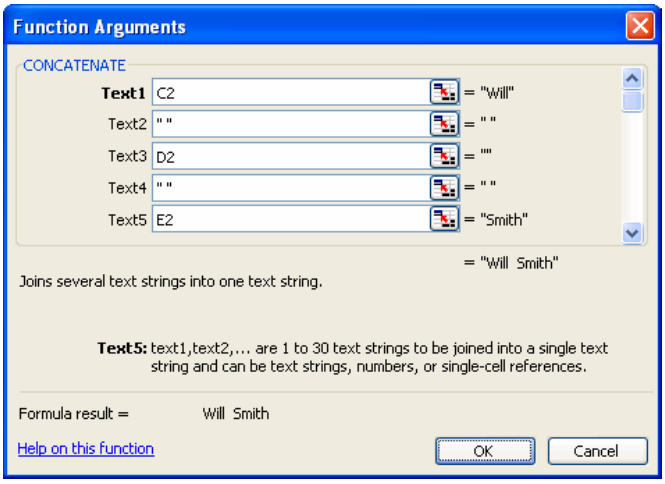
WEEKDAY – Returns a number from 1 to 7 identifying the day of the week of a date.
Insert a column to the right of PaymentDate
Click the Insert Function button
Search for **WEEKDAY**
Serial_number – the number that represents a date
Return_type – 1 (Sun. = 1)
2 (Mon. = 1)

MOD – returns the remainder after a number is divided (we use this to identify even dollar transactions)
Insert a column to the right of PaymentAmt
Click the Insert Function button
Search for **MOD**
Number – Dollar amount
Divisor – The number by which you want to divide by (10)



SPLIT – Split the contents of cells across multiple cells using the Text to Columns tool
Open the EmployeeInfo spreadsheet
Insert three columns to the right of FullName
Copy the contents of FullName into the third empty column you created (when you split the cells, the FullName column will become the last name
Highlight the original FullName column
Data → Text to Columns
Click next for Delimited
Check only the Comma box to split the last name from the first name and middle initial
Change the format to Text and click Finish
Highlight the column that now contains the first name and middle initial and repeat the process. This time select the Space check box. For the first column, click the radio button to NOT import it.

CONCATENATE – Joins several text strings into one
Rearrange the columns (First Name, MI, Last Name)
Insert a column to the right of the Last Name
Click the Insert Function button
Search for Concatenate
Enter the cells in the Text Fields. After each Cell, enter " " to allow for a space between the names.



You will notice that the names that do not have a middle initial will have an extra space between the First and Last Names. To fix this, you can use the TRIM function which removes all the blank spaces except for those between words. If you click on a cell, the formula should read **=CONCATENATE(C2," ",D2," ",E2)**
In the formula bar, change the formula to **=TRIM(CONCATENATE(C2," ",D2," ",E2))**. This should remove the extra spaces.

Duplicate Payments

Used to determine if multiple records have the same value in specified fields.

Sort the table by PayeeID, PaymentAmt, and Invoice
Create a new column called Duplicate
In the second record, enter the formula **=IF(AND(E7=E6,P7=P6,T7=T6),"Dup","Not a Dup")**
Copy the formula down

Gaps

Determines if an item is missing from a list

Open DOESVendor file
Sort table by VendorID
Create a column called Gaps
In the second record, enter the formula **=IF(A3-A2<>1,(A3-A2)-1,0)**
If there are no gaps, the result is 0. If there are gaps, the formula will calculate how many items in the list are missing.

Benford's Analysis

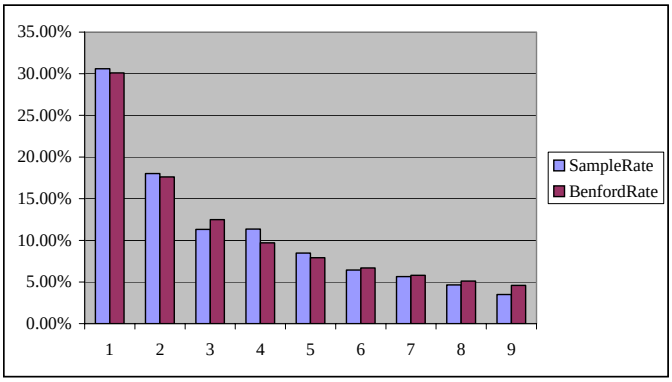
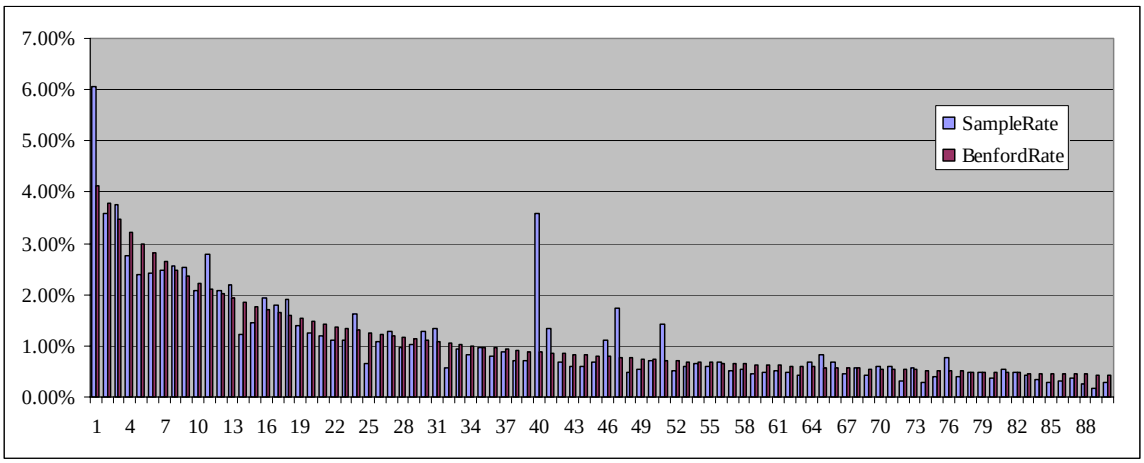
Insert a column to the right of PaymentAmt (PaymentAmt for Benford)
To remove the decimal, divide the PaymentAmt field by .01 (e.g., P6/.01)
Create two additional columns to the right of PaymentAmt for Benford (1st digit & 2 digit)
Use the **LEFT** function to grab the first digit and first 2 digits of the PaymentAmt for Benford field =Left(Q6) =Left(Q6,2)

Insert a new worksheet called Benford
In column A, enter the numbers 1-9
In column B, use the COUNTIF function to count the number of instances the numbers in Column A occur =COUNTIF(Agy42000.B0403.E0603!$R$6:$R$4453,Benford!A2)

In Column C, calculate the Sample Rate (the percentage that number occurs) =B2/$B$11
In Column D, calculate the Benford Rate (the percentage that number should occur) =LOG10(1/A2+1)

Select the Sample Rate and Benford Rate columns and chart the results.

Number	Count	SampleRate	BenfordRate
1	1361	30.60%	30.10%
2	802	18.03%	17.61%
3	503	11.31%	12.49%
4	505	11.35%	9.69%
5	377	8.48%	7.92%
6	286	6.43%	6.69%
7	252	5.67%	5.80%
8	207	4.65%	5.12%
9	155	3.48%	4.58%
	4448		



Repeat the results for the first 2 digits.