Abnormal Transactions Testing with Benford's Law

Lab Insight: An important role of the internal auditor is to detect fraudulent and suspicious transactions. While traditional audit approaches randomly select samples from daily transactions, data analytics provides multiple approaches to achieve data reduction and to transfer auditor's focus to more critical accounts that require further investigation. In this lab, we use Benford's Law to set a boundary of normal behaviors and narrow down the investigation to abnormal transactions.

Ask the Question: Could Benford's Law help internal auditors identify suspicious activities?

Master the Data: Luckyme is a cloth store. Luckyme begins business on 1/1/2024 and has 200 sales in the first month. The sales are recorded in the Excel spreadsheet "Abnormal Transactions Testing.xlsx".

Data Dictionary:

> **Order_Number:** The number on the receipt
> **Sales:** Sales revenue of each transaction in dollars

Part 1:

Perform the Analysis:

1. Open Excel file "**Abnormal Transactions Testing.xlsx**".

2. Browse the spreadsheet very quickly to make sure there aren't any obvious errors in the Excel file.

3. Create a new worksheet which is a copy of the Sales worksheet. Name the new worksheet **FirstDigit**.

4. In cell C1 of the Sales worksheet, type First_Digit. In cell C2, enter the formula "=Left(B2,1)" and copy the formula down the rest of the column.

	A	B	C	D	E	F	G	H	I
1	Order_Number	Sales	First_Digit						
2	00001	129.33	=LEFT(B2,1)						
3	00002	234.85	2						
4	00003	934.83	9						
5	00004	485.48	4						
6	00005	384.58	3						
7	00006	173.65	1						
8	00007	283.99	2						
9	00008	113.98	1						

5. Create a new worksheet called **BenfordLaw**.
   a. In cell A1, type **First_Digit.**
   b. In cell A2, type the number "1".
   c. In cell A3, type the number "2".
   d. Select A2 and A3 and drag down to A10 to extend the series (so the column should be from 1 to 9).

6. In B1, label the column Count.   In B2, insert the COUNTIF function.   For the range, highlight the entire column C in the FirstD worksheet.   For the criteria, enter A2.   Copy the formula down.

	A	B	C	D	E
1		Count			
2		1 =COUNTIF(FirstD!C:C,A2)			
3	2	42			
4	3	27			
5	4	20			
6	5	16			
7	6	21			
8	7	10			
9	8	10			
10	9	8			
11					
12					

7. In cell B11, use AutoSum to determine the total number of transactions in cells B2 to B10 (hint select all the cells under the header of **Count** > go to **Home** > Click **AutoSum** in the section of Editing. Alt + = is also a keyboard shortcut to help you AutoSum.)

Lab Benford's Law - Excel

File   Home   Insert   Page Layout   Formulas   Data   Review   View   Pov

SUM   ✕ ✓ *fx*   =SUM(B2:B10)

	A	B	C	D	E	F	G
1	First_Digit	Count					
2		1	46					
3		2	42					
4		3	27					
5		4	20					
6		5	16					
7		6	21					
8		7	10					
9		8	10					
10		9	8					
11		=SUM(B2:B10)					

8. The next step is to calculate the percentage of each digit and compare it with expected percentage of Benford's Law. In cell C1, type **Percentage.**   In cell C2, then *divide* the count of first-digit numbers by the total number of transactions in B11. The *formula* to calculate the percentage of first-digit 1 is as noted below. Once this is done, *copy* this formula all the way down to C10.

9. In cell D1, *type* "Benford's Law".   In D2, *calculate* the expected frequency based on Benford's Law using the formula "**=LOG10(1/A2+1)**".

Copy the formula to the last row of **First_Digit.**

10. Format the Percentage and Benford's Law columns as percentage.   (Hint: Select columns of Percentage and Benford's Law > Go to **Home > Number > Percentage**)
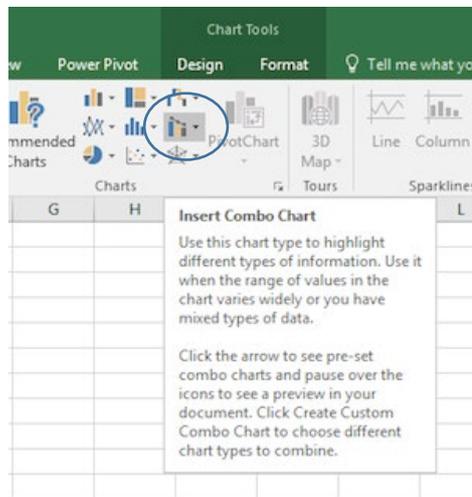

11. Calculate the difference next between the expected percentage of first digits

based on Benford's Law and the percentage of our dataset as follows (after copying down the column):

	A	B	C	D	E
1	First_Digit	Count	Percentage	Benford's Law	Difference
2	1	46	23%	30%	=C2-D2
3	2	42	21%	18%	3%
4	3	27	14%	12%	1%
5	4	20	10%	10%	0%
6	5	16	8%	8%	0%
7	6	21	11%	7%	4%
8	7	10	5%	6%	-1%
9	8	10	5%	5%	0%
0	9	8	4%	5%	-1%
1		200			

12. Insert the Combo Chart
    a. Highlight the Percentage and Benford's Law columns including the headers.
    b. Go to Insert > Charts > Insert Combo Chart. Change the name of the Chart to Benford's Law.

	A	B	C	D	E	F
1	First Digit	Count	Percentage	Benford's	Difference	
2	1	46	23%	30%	-7%	
3	2	42	21%	18%	3%	
4	3	27	14%	12%	1%	
5	4	20	10%	10%	0%	
6	5	16	8%	8%	0%	
7	6	21	11%	7%	4%	
8	7	10	5%	6%	-1%	
9	8	10	5%	5%	0%	
10	9	8	4%	5%	-1%	
11		200				
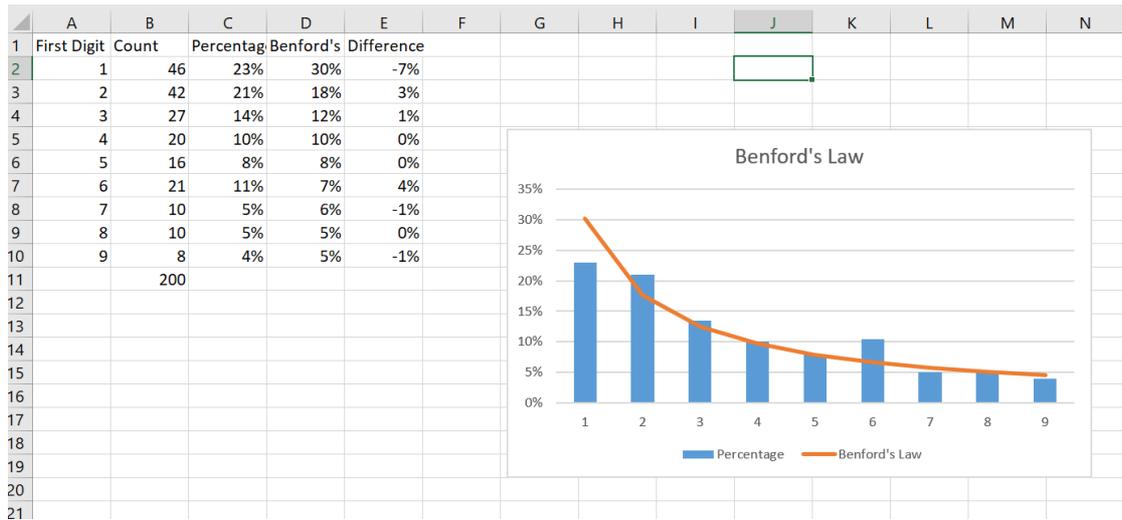
Now, we can compare the distribution of each first-digit in our dataset with the expected values based on Benford's Law. Transaction counts deviating from the expected count (per Benford's Law) may indicate fraud and require internal audit investigation.

The above graph suggests that fewer than expected sales transactions started with the number 1, but sales transactions starting with the number 2 and 6 occurred more frequently than expected. Internal auditors should focus their attention on the sales beginning with 2 and 6 as sales beginning with 1 may have been altered to start with 2 and 6.

13. We can also use a statistical test named chi-square to test if the distribution of first-digits of our dataset is significantly different than that expected by Benford's Law.

We have two opposing hypotheses:
H_0: Our dataset's first digits conform to Benford's Law.
H_1: Our dataset's first digits <u>do not</u> conform to Benford's Law.

When the null hypothesis is rejected (p-value <=0.05), additional investigations are needed to test for fraud or irregularities.

Add an additional column called "Expected Count". Expected Count that equals D2* B11 or the Benford's percentage multiplied by the total amount of transactions. Copy down the column. (Note: You may need to move the chart over to the right. In addition, round to the nearest whole number.)

CHISQ.TEST | =D2*B11

	First Digit	Count		Percentage	Benford's Law	Difference	Expected Count	
	A	B		C	D	E	F	G
1	First Digit	Count		Percentage	Benford's Law	Difference	Expected Count	
2	1		46	23%	30%	-7%	=D2*B11	
3	2		42	21%	18%	3%	35	
4	3		27	14%	12%	1%	25	
5	4		20	10%	10%	0%	19	
6	5		16	8%	8%	0%	16	
7	6		21	11%	7%	4%	13	
8	7		10	5%	6%	-1%	12	
9	8		10	5%	5%	0%	10	
10	9		8	4%	5%	-1%	9	
11			200					
12								
13								
14								

Use the Chi-square test function to determine the p-value. If the p-value of Chi-square is less than or equal to 0.05, then we reject the null hypothesis that the distributions are not different from each other and investigate the reasons the distributions are different from each.

C13 | =CHISQ.TEST(B2:B10,F2:F10)

	A	B	C	D	E	F	G
1	First Digit	Count	Percentage	Benford's Law	Difference	Expected Count	
2	1	46	23%	30%	-7%	60	
3	2	42	21%	18%	3%	35	
4	3	27	14%	12%	1%	25	
5	4	20	10%	10%	0%	19	
6	5	16	8%	8%	0%	16	
7	6	21	11%	7%	4%	13	
8	7	10	5%	6%	-1%	12	
9	8	10	5%	5%	0%	10	
10	9	8	4%	5%	-1%	9	
11		200					
12							
13		chi square test	=CHISQ.TEST(B2:B10,F2:F10)				
14							
15							
16							

First Digit	Count	Percentage	Benford's Law	Difference	Expected Count
1	46	23%	30%	-7%	60
2	42	21%	18%	3%	35
3	27	14%	12%	1%	25
4	20	10%	10%	0%	19
5	16	8%	8%	0%	16
6	21	11%	7%	4%	13
7	10	5%	6%	-1%	12
8	10	5%	5%	0%	10
9	8	4%	5%	-1%	9
	200				
chi square test		0.30			

The p-value for the Chi-square statistic is 0.30.   So, we do not reject H_0, indicating that while the actual first digit distribution is not exactly the same as the expected distribution of Benford's Law, the differences are not significant.

Share the Story:

Benford's Law can help internal auditors identify frauds and errors. Benford's Law helps identify specific accounts or transactions that require additional verification, increasing work efficiency and effectiveness (versus random sampling).