**Benford’ Law Or the first Digit Law**

It’s an observation about the frequency distribution of the leading digits in sets of numerical data. In naturally occurring collections of numbers you leading digit is likely to be small. So, if your data obeys the law you are going to find out that number 1 appears as the leading digit 30% of the time all the way down to the number nine as the leading digit that’s going to occur less than five percent of the time.

**Why Does Bedford’s law matter?**

When fraudsters are fabricating data they may not know to create their fake data in a manner that conforms to Benford’s law. You can run this analysis to detect fake data or at least doubt about its authenticity.

Steps

Create a calculate field **Leftmost Integer** to change the sales to string and derive the first digit that is the leftmost digit.

**LEFT(STR([Sales]),1)**

Drag the new calculated field to the rows shelf.

Since Benford’s law doesn’t Zeros in the analysis exclude the Zero.

Drag the number of records to the Rows shelf and change the table calculation to percentage of total.

You can see number 1 as the leftmost integer occurs 30.87% o f the time all the way down to nine as the leading digit occurs almost 4.38% of the time.

To verify how does Benford’s law apply here create a Calculate field

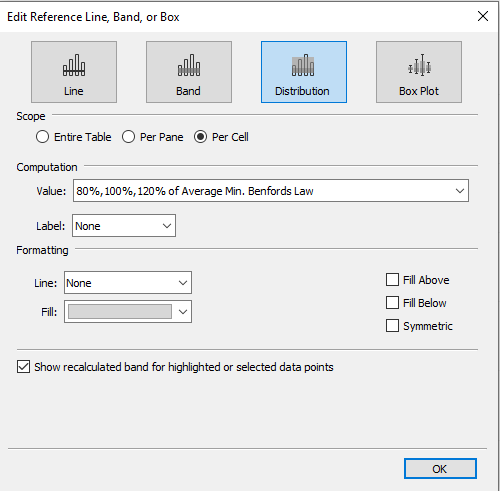
**Benfords Law**

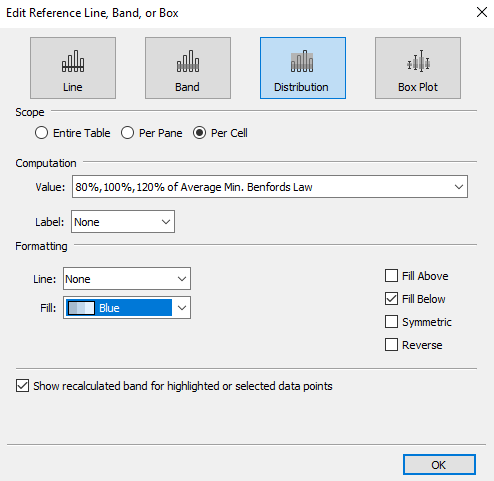
**LOG(INT([Leftmost Integer])+1)-LOG(INT([Leftmost Integer]))**

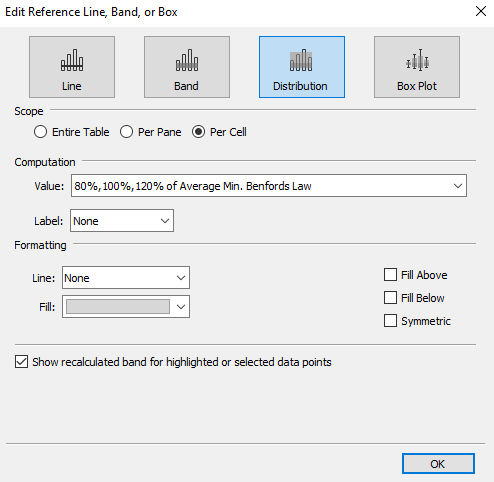
Drag the Benford law to details and Change the measure to minimum.

Go to the analytics tab and Drag the Distribution band to the cell

Edit the reference lines as follows







Change the view to entire view and format the min(Benfords law) in the Details shelf to percentage.

In order to Highlight the differences, between our calculation and Benford’s Law create a calculated field

**Pct Diff Benford's Law**

**SUM([Number of Records]) / TOTAL(SUM([Number of Records]))-MIN([Benfords Law])**

**Or**

**COUNT([Orders]) / TOTAL(COUNT([Orders]))-MIN([Benfords Law])**

Drag the calculated field to text and change the format to percentage.

Verify the differences if the data is ok.