A Data Analysis Based Framework to Detect Anomalies in Large Data Sets Using Benford's Law

Mustafa Canbolat and D. Donald Kent, Jr

The *BRC Academy Journal of Business* 7, no. 1 (2017): 99–110

http://dx.doi.org/10.15239/j.brcacadjb.2017.07.01.ja05

Web Appendix

http://dx.doi.org/10.15239/j.brcacadjb.2017.07.01.wa05

Figure 1: An Overview of Fraud Detection Methods

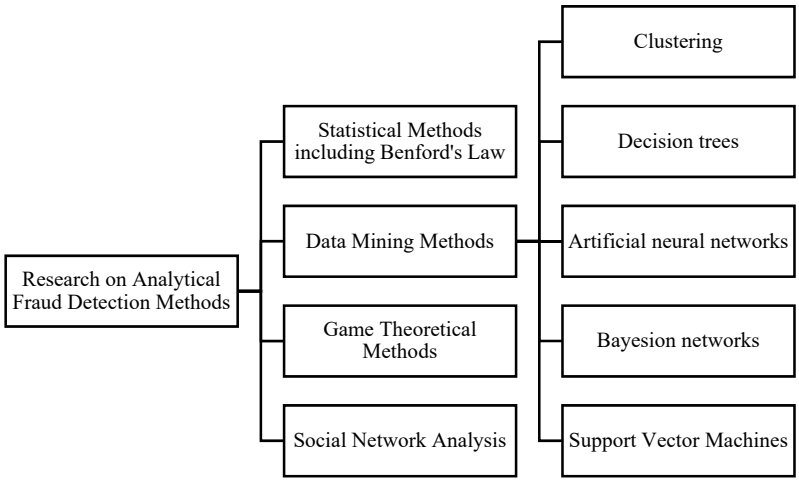

Figure 2: Data Analysis Framework

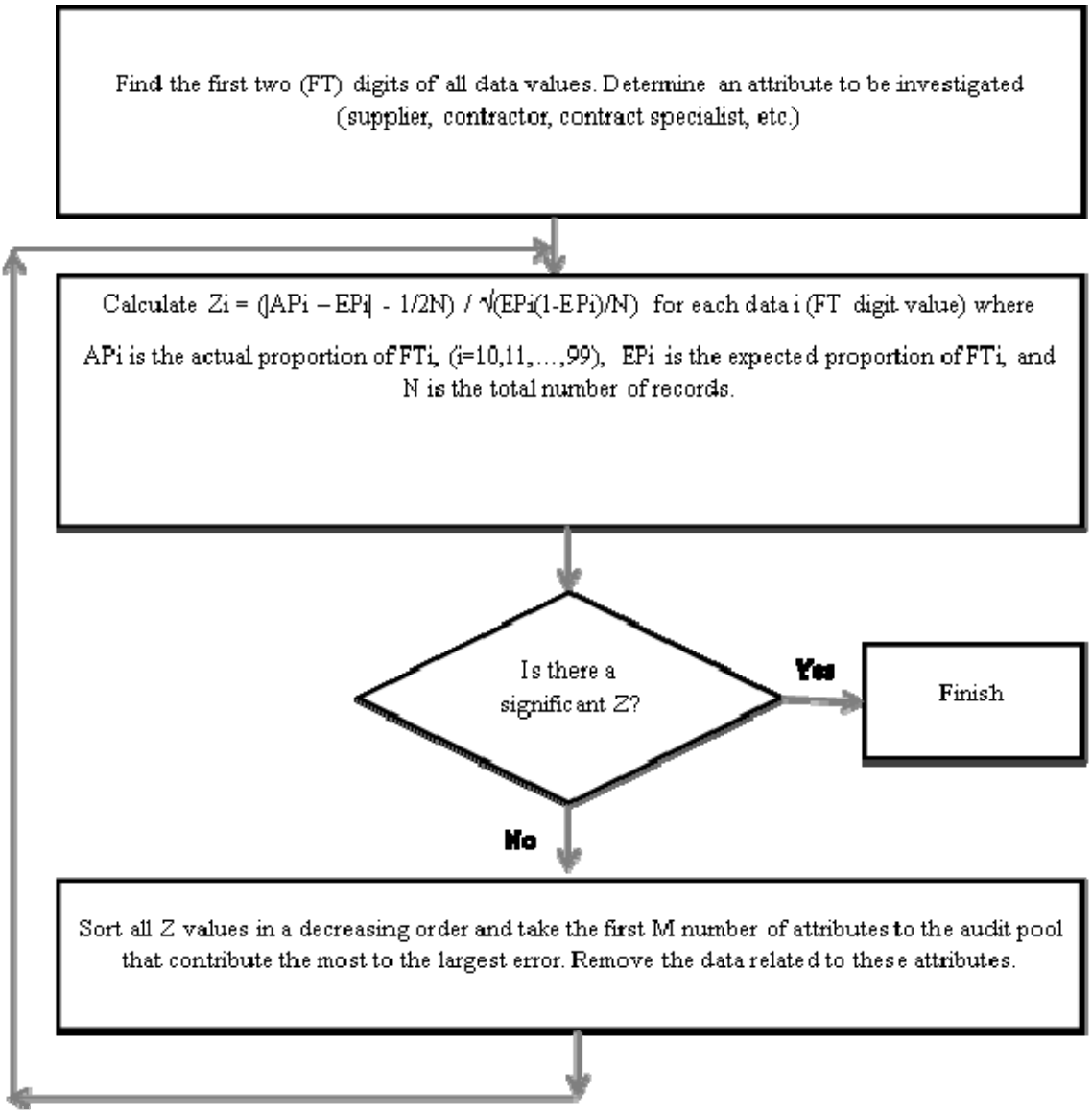

Figure 3: First two digit distribution of the actual and the expected proportions

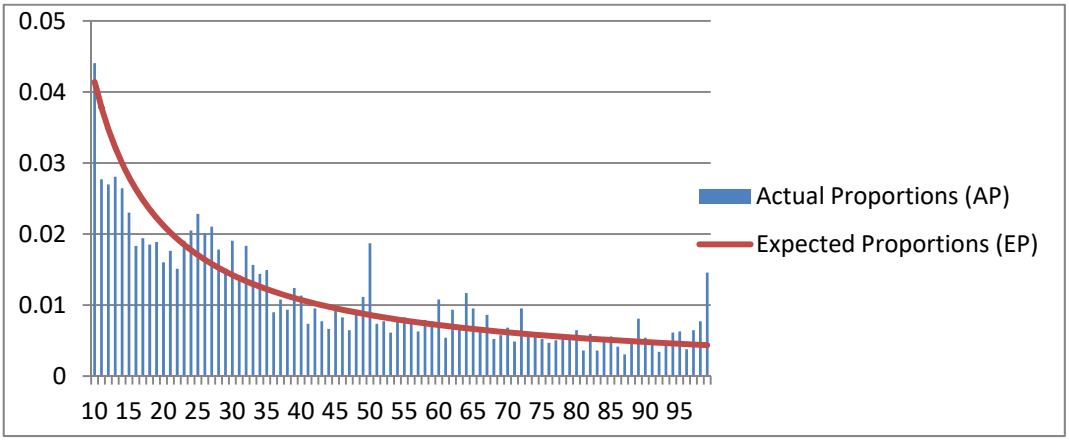

Table 1. Results before the First Iteration

First Two Digits	Frequency	Observed %	Expected %	Z-value
99	81	1.46%	0.44%	11.44
50	104	1.87%	0.86%	8.09
64	65	1.17%	0.67%	4.44
11	154	2.77%	3.78%	3.91
16	102	1.83%	2.63%	3.68
98	43	0.77%	0.44%	3.64
89	45	0.81%	0.49%	3.38

Table 2. Results before the Second Iteration

First Two Digits	Frequency	Observed %	Expected %	Z-value
50	94	1.88%	0.86%	7.71
99	49	0.98%	0.44%	5.71
25	126	2.51%	1.70%	4.39
27	117	2.34%	1.58%	4.23
32	99	1.98%	1.34%	3.88
72	51	1.02%	0.60%	3.75
11	140	2.79%	3.78%	3.62