In the Trenches with Tableau, Pivot Tables, Fuzzy Lookup & Benford's Analysis

Billy Morehead, Ph.D., CPA, CGFM
Professor of Accountancy & Director of the MAcc
AGA Past National President
Mississippi College
April 2022

1

AI & ML Usage: Not Slowing Down

Accelerating AI Adoption

Chatbots			
	Cities	Counties	States
Not using; plan to implement in 12-18 mo.	37%	34%	17%
In use	15%	15%	44%
In use; plan to upgrade in 12-18 mo.	13%	25%	39%
AI/ML for Cybersecurity			
Not using; plan to implement in 12-18 mo.	23%	20%	28%
In use	23%	32%	28%
In use; plan to upgrade in 12-18 mo.	26%	29%	25%
Business Intelligence/Data Analytics			
Not using; plan to implement in 12-18 mo.	10%	7%	2%
In use	34%	33%	37%
In use; plan to upgrade in 12-18 mo.	55%	53%	59%

Source: CDG Digital Cities, Counties and States Surveys

Source: Government Technology - 2021

2

Innovations in Government

AI, BI & ML - COVID 19 Accelerated
- National Trend
- Broad Development Challenges (broadband, data access, recidivisim, need)
- Poised for Growth

PEOPLE CHALLENGES:
- Improve service delivery (user experience)
- Institutional knowledge (gaps in subject matter expertise
- Effective, efficient, streamlined processes
- Collaboration & productivity across teams

3

Government Technologies - 2021

3

Innovations in Government

PREPARE FOR SUCCESS:
- Evolving & emerging technologies
- Legacy processes (and people)
- Data driven & automated organizations
- Enterprise data cloud platforms help
- Data pipeline
- Security and Governance
- Cultural changes (society, job changes, ADA)
- Right skillsets in staff (IT, data scientists, analysts
- Organizational maturity
- Look for quick wins

4
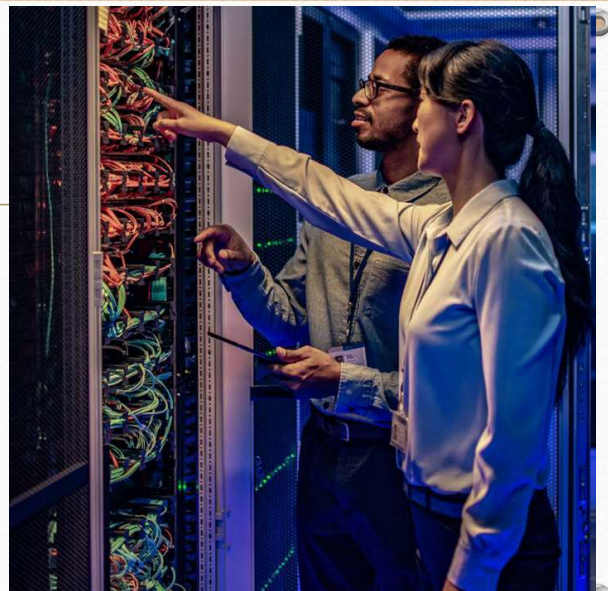
Government Technologies, 2021 - Photo: e.Republic. 2020

4

Data Analytics – Defined

- The process of evaluating data with a purpose of drawing conclusions to address business questions. Effective Data Analytics provides a way to search through large structured and unstructured data to discover unknown patterns or relationships.

- In other words, Data Analytics often involves the technologies, systems, practices, methodologies, databases, statistics, and applications used to analyze diverse business data to give organizations the information they need to make sound and timely business decisions…transforming raw data into knowledge to create value.

5

Source: Data Analytics for Accounting, 2nd Ed. Richardson, Teeter, & Terrell. 2021

5

How Data Analytics Affects Business

- Generate up to $5 Trillion in value per year

- Transform the how companies run their businesses

- Discover various buying patterns of customers

- Investigate anomalies

- Forecast future possibilities

- Execute more directed marketing campaigns

- Give a competitive advantage over others not using DA

- Better identification of risks and opportunities

- Improved internal processes, productivity, utilization and growth

6

Source: Data Analytics for Accounting, 2nd Ed. Richardson, Teeter, & Terrell. 2021

6

How Data Analytics Affects Accounting

AUDIT:

- Audit MUST better embrace technology.
- Technology will enhance the quality, transparency, and accuracy of the audit.

TAX:

- Sophisticated Tax Planning (minimize tax liability, avoid or prepare for audits, predictive analytics)

FINANCIAL REPORTING:

- Improve quality of estimates and valuations
- Address A/R Collection
- Allowance for losses
- Inventory obsolescence (out of date? need to discount and sale?)
- Valuation at Cost or Market
- Goodwill been impaired?
- Valuing Warranty Claims or Litigation

Source: Data Analytics for Accounting, 2nd Ed. Richardson, Teeter, & Terrell. 2021

7

Data Analytics Process – IMPACT Cycle
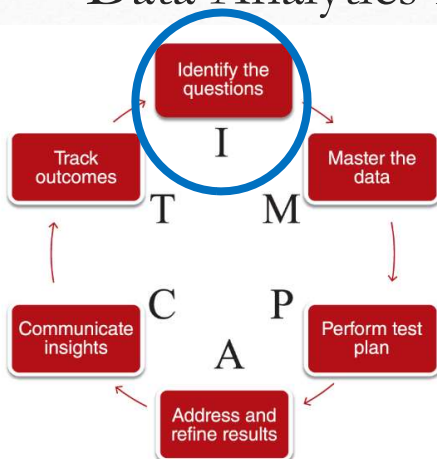
Source: Isson, J. P., and J. S. Harriott. *Win with Advanced Business Analytics: Creating Business Value from Your Data.* Hoboken, NJ: Wiley, 2013.

Identify the Questions:

- Are employees circumventing internal controls?
- Any suspicious travel and entertainment expenses?
- How can we increase the "add-on" sales of items to customers?
- Are customers paying timely?
- Finding risky transactions?
- Who authorizes checks over $100,000?
- How can errors be identified?

Source: Data Analytics for Accounting, 2nd Ed. Richardson, Teeter, & Terrell. 2021

8

Data Analytics Process – IMPACT Cycle
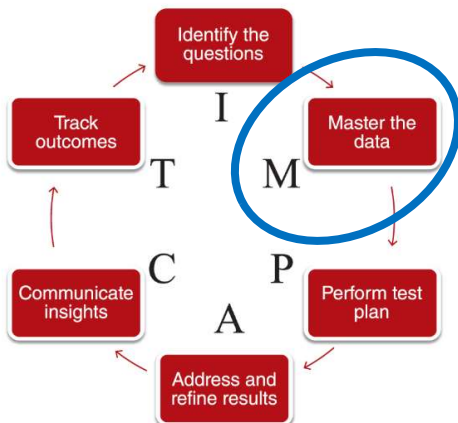
Master the Data:

- Review data availability in internal systems
- Review data availability in external networks
- Use of data dictionaries and other contextual data
- Extraction, transformation and loading (ETL)
- Data validation and completeness (reliability)
- Data normalization (redundancy, integrity)
- Data preparation and scrubbing (50 to 90% of DA time spent cleaning data to be analyzed)

Source: Isson, J. P., and J. S. Harriott. *Win with Advanced Business Analytics: Creating Business Value from Your Data.* Hoboken, NJ: Wiley, 2013.

Source: Data Analytics for Accounting, 2nd Ed. Richardson, Teeter, & Terrell. 2021
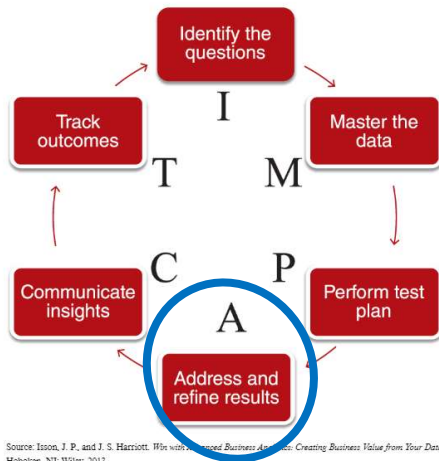
9

9

Data Analytics Process – IMPACT Cycle
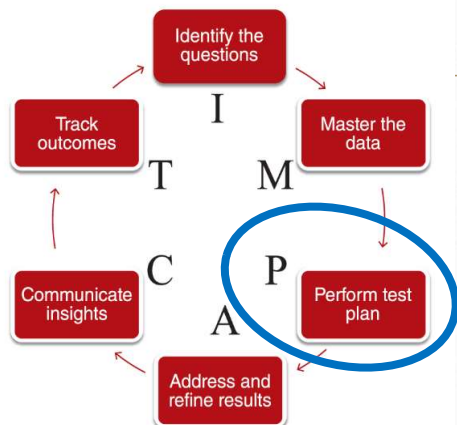
Perform the Test:

- Using all data available, see if we can identify relationships between the response (dependent) variables and those items that affect all response (predictor, explanatory, independent) variables.
- **Classification** – assign units to a few categories (teacher/student)
- **Regression** – predict specific dependent variables based on independent variable inputs (loan default – income, GPA, age)
- **Similarity Matching** – match based on known data
- **Clustering** – Segment individuals into groups
- **Co-occurrence grouping** – Associations between individuals and transactions; "frequently bought together"
- **Profiling** – Characterizing "typical" behavior
- **Link Prediction** – Predict a relationship between two data items
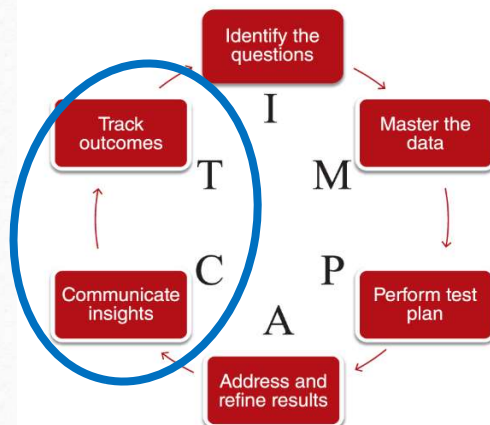- **Data Reduction** – reduce volume (highest cost, risk, impact)

Source: Isson, J. P., and J. S. Harriott. *Win with Advanced Business Analytics: Creating Business Value from Your Data.* Hoboken, NJ: Wiley, 2013.

Source: Data Analytics for Accounting, 2nd Ed. Richardson, Teeter, & Terrell. 2021

10

10

Data Analytics Process – IMPACT Cycle

Address and Refine Results:

- DA is iterative
- Slice and Dice data
- Find correlations
- Ask additional questions
  - Of ourselves
  - Of others (colleagues, stakeholders, etc.)
- Refine, revise and rerun the analysis

Source: Isson, J. P., and J. S. Harriott. *Win with Advanced Business Analytics: Creating Business Value from Your Data.* Hoboken, NJ: Wiley, 2013.

11

Source: Data Analytics for Accounting, 2nd Ed. Richardson, Teeter, & Terrell. 2021

11

Data Analytics Process – IMPACT Cycle

Communicate Insights & Track Outcomes:

- Static Reports
- Digital Dashboards
- Data Visualizations

Source: Isson, J. P., and J. S. Harriott. *Win with Advanced Business Analytics: Creating Business Value from Your Data.* Hoboken, NJ: Wiley, 2013.

12

Source: Data Analytics for Accounting, 2nd Ed. Richardson, Teeter, & Terrell. 2021

12

Let's Dig Into Tableau –Procurement Card Data
(Go to Participant Handout and Live Tableau Demonstration)

- Pcard Data Set from State of Oklahoma Website:

- https://data.ok.gov/dataset/state-of-oklahoma-purchase-card-pcard-fy-2021/resource/520e34b4-8fd9-49ca-9081-25a5a1a0aa43

13

13

Let's Dig Into Tableau –Procurement Card Data
(Go to Participant Handout and Live Tableau Demonstration)

14

Pivot Tables

- Allow you to quicly summarize large amounts ot data. In Excel, click Insert>PivotTable, choose your data source, then click the checkmark next to or drag your fields to the appropriate boxes in the PivotTables Fields pane to identify filters, columns, rows or values.

- Rows – show the main item of interest. You usually want master date here…such as customers, products or accounts.

- Columns – Slice the data into categories or buckets. Most commonly, columns are used for time (years, quarters, months or dates).

15

15

Pivot Tables

- Values – This area represents the meat of your data. Any measure that you would like to count, sum, average or otherwise aggregate should be placed here. The aggregated values will combine all records that match a give row and column.

- Filters – placing a field in the Filters area will allow you to filter the data based on that field, but it will not show that field in the data. For example, if you wanted to filter based on a date, but didn't care to view a particular date, you could use this area of the field list.

16

16

PivotTable Example – Live
(Refer to Your Handout)
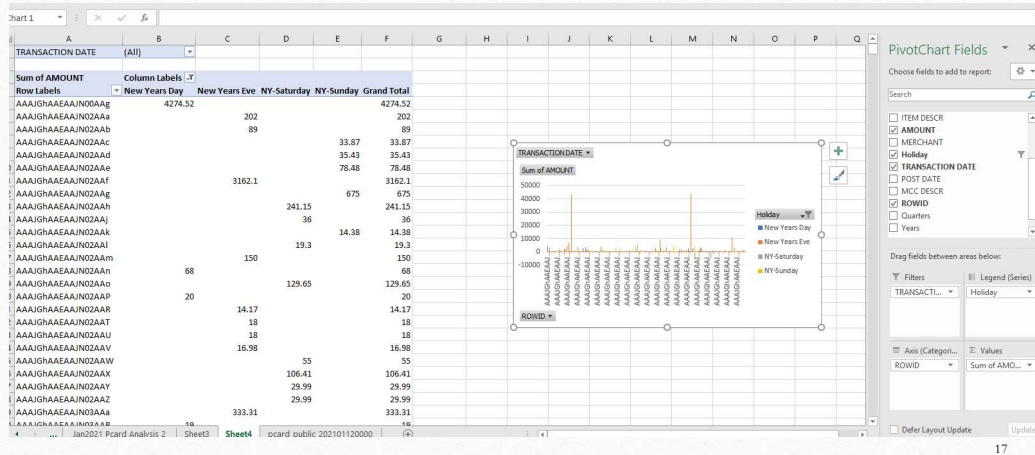
17

Fuzzy Lookup

- Finding similar items with a less than 100% match

- Example: "456 Third Avenue" and "456 Third Av.e." are not the exact same value but might be the same location.

- Determine a threshold (ex: 50%)

- Do you want Matching or Non-Matching?

18

Fuzzy Lookup

- Inner Join – shows ONLY records from both tables that match and excludes items that don't match.

- Left Join – shows all records from Table 1 and only matching records in Table 2.

- Right Join – shows all records from Table 2 and only matching records from Table 1.

- Outer Join – shows all records from both tables – matching and non-matching

19

19

# LiveExample – See handout

20

20

Example Results

Supplier_Contact	Supplier_Billing_Add	Supplier_Billing	Supplier_	Supplier_	Supplier_Billing_	Supplier_	Supplier_Physical_S
1 Adrienne Yates	377 High Ridge Ct	Hamtramck	MI	48212	8501852845	n30	377 High Ridge Ct
1 Kyla Stout	447 Riverview Ave.	Upper Darby	PA	19082	3392868017	2/10n30	447 Riverview Ave.

21
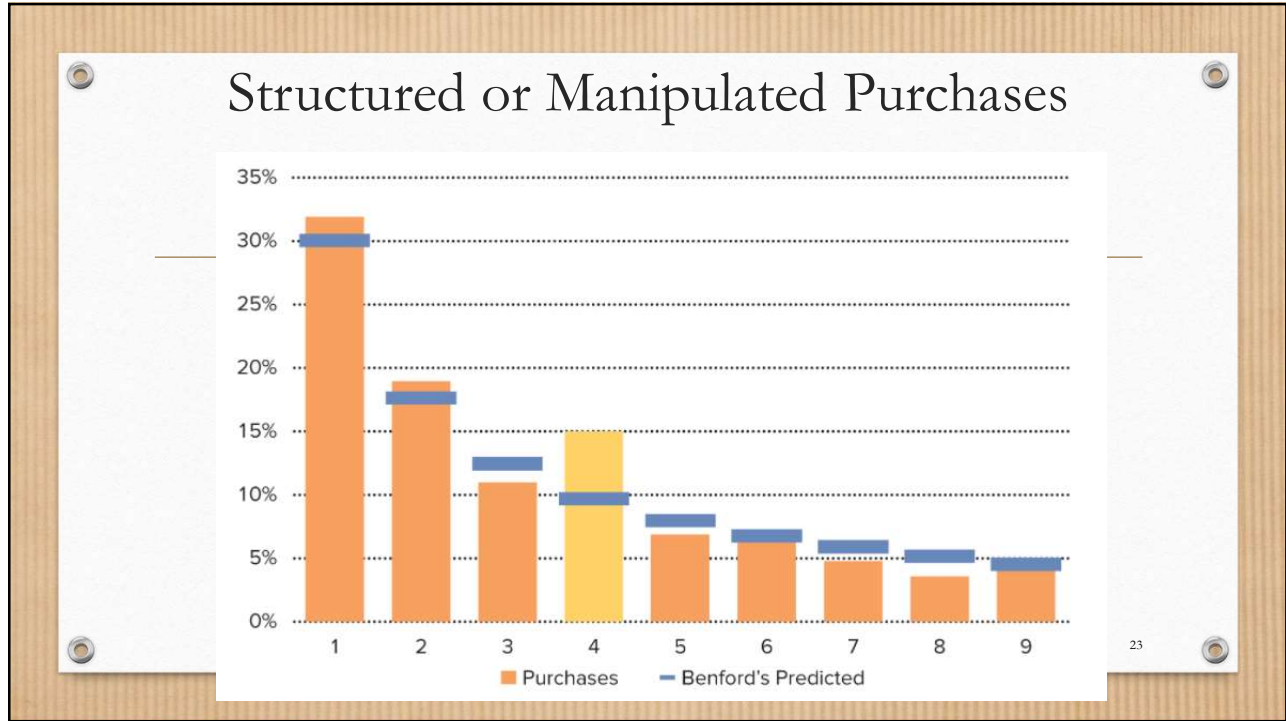
Benford's Law

21

22

Structured or Manipulated Purchases

23
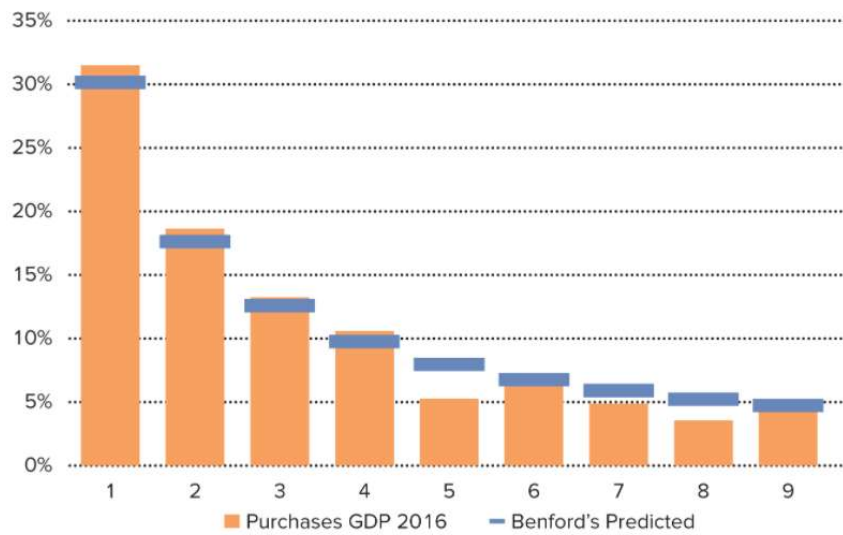
23

Let's Create Our Own Benford's Analysis Chart
(See Participant Handout)

Digit	Actual Count	Actual %	Expected %
1	=COUNTIF([Range],[Digit])	(=[Actual Count]/SUM[Actual Count])	30.1%
2	. . .	. . .	17.6%
3	. . .	. . .	12.5%
4	. . .	. . .	9.6%
5	. . .	. . .	7.9%
6	. . .	. . .	6.7%
7	. . .	. . .	5.8%
8	. . .	. . .	5.1%
9	. . .	. . .	4.6%
	=SUM([Actual Count])	=SUM([Actual %])	=SUM([Expected %])

24

24

Leading Digit	Benford's Frequency Distribution Chart			
-				
-	Digit	Actual Count	Actual %	Expected %
-	1	7394	30.068	30.1
9	2	4543	18.474	17.6
-	3	3003	12.212	12.5
1	4	2361	9.601	9.6
1	5	1878	7.637	7.9
-	6	1522	6.189	6.7
3	7	1404	5.709	5.8
-	8	1178	4.790	5.1
3	9	1308	5.319	4.6
-		24591	100.00	99.9
5				

25

Benford's Analysis of January 2021 PCard Data
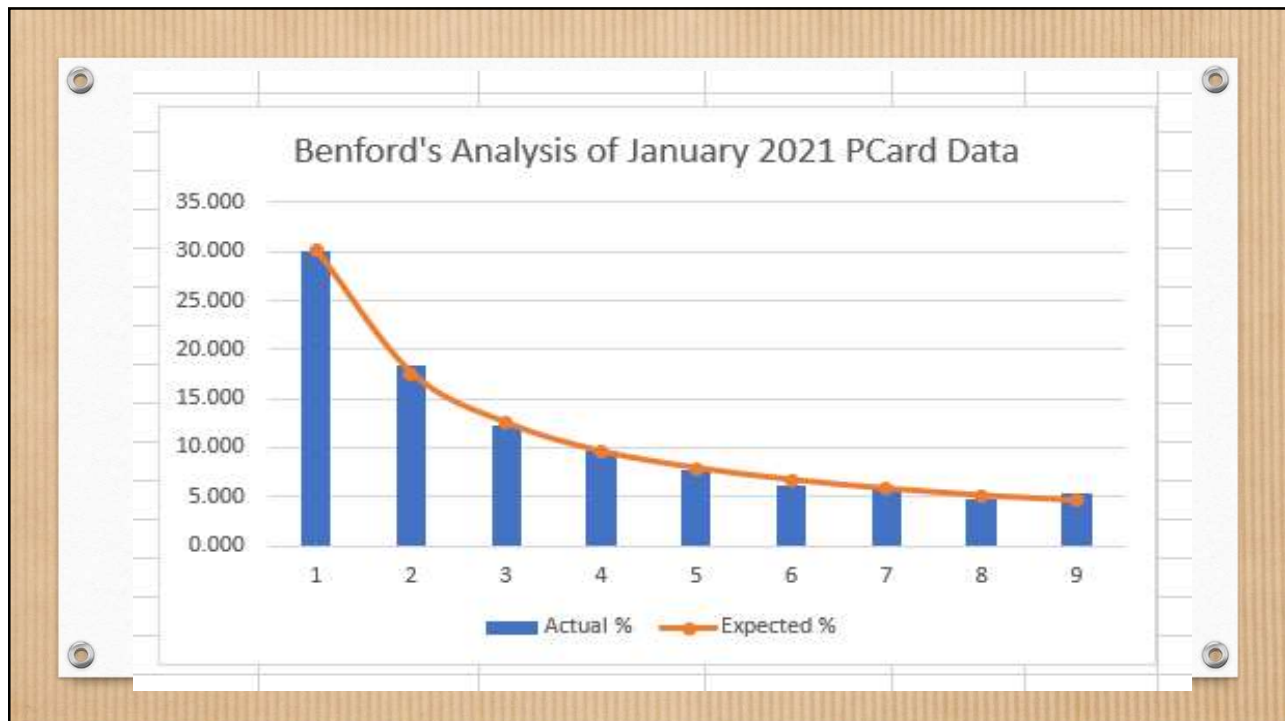
26

In the Trenches – WOW!

- Experiment
- Undo Button
- Play with your data
- Get creative
- BUT, ABOVE ALL…DON'T FORGET TO SAVE YOUR ORIGINAL DATA SET!

27

27

Contact Information:

Billy Morehead

Professor of Accountancy

Mississippi College

morehead@mc.edu

601-925-7742

28

28