Accounting Analytics

Tom Adams, Ph.D., CPA, CGMA
May 23, 2025

Tom Adams, Ph.D., CPA, CGMA

Associate Professor of Accounting at La Salle University

Tom Adams joined the Accounting Department in the School of Business in 2019. Previously, he was an assistant professor at the University of Connecticut (2016-2019). He completed his Ph.D. in accounting at Temple University (2011-2016).

He is a proud La Salle University graduate. He earned his Bachelor of Science in Business Administration (accounting and finance) at La Salle in 2005. After graduation, he worked as a Senior Associate in the Audit Practice of KPMG (Philadelphia office; 2005-2009) and as a Senior Financial Analyst in the Investor Relations department of Teleflex Incorporated (NYSE: TFX; 2009-2011).

Tom is originally from Northeast Philadelphia and attended St. Bernard's Grade School and Father Judge High School. He is currently tea*ching Accounting Analytics* (both undergraduate and graduate), *Accounting Information Systems, Advanced Auditing/Financial Forensics, Auditing & Assurance Services,* and *Introduction to Financial Accounting*. His research has been published in *Current Issues in Auditing*, the *Journal of Accounting, Auditing and Finance*, the *Journal of Accounting and Public Policy*, the *Journal of Futures Markets*, *Review of Financial Economics*, *CPA Now - PICPA's Blog for Pennsylvania CPAs*, *Estate Planning*, the *Journal of Accountancy*, *Practical Tax Strategies*, and the *Tax Insider*.

La Salle is a place near and dear to Tom's heart. His grandfather and father also attended the university and he considers La Salle home.

What is going on the world of analytics and tech?

Date	Topic	Image	Article Title
Dec 6, 2024	Analytics		Appeals Court Upholds U.S. Ban of TikTok
Dec 5, 2024	Analytics		Bitcoin Hits $100,000, Lifted by Hopes of a Crypto-Friendly Washington
Dec 5, 2024	Analytics		Robinhood's Hot New Tech Bet: The Desktop Computer
Dec 3, 2024	Analytics		AI and the Automation of Work for Gen Z
Dec 3, 2024	Analytics		Why Hackers Want Your Health Information
Dec 3, 2024	Analytics		Microsoft Faces $1.27 Billion Antitrust Lawsuit Over Cloud Services in U.K.
Dec 3, 2024	Analytics		Cyberattack and Financial Troubles Force Stoli's U.S. Arm to File for Bankruptcy
Dec 2, 2024	Analytics		Do Your Passwords Meet the Proposed New Federal Guidelines?
Nov 30, 2024	Analytics		The iPhone's 'Lockdown Mode': What It Is and Who Should Consider Using It
Nov 29, 2024	Analytics		For EV Startups, Things Are Going From Bad to Worse

Articles

1. Data analytics for accounting and identifying questions

1.1. Define data analytics

Data Analytics is the process of evaluating data with the purpose of drawing conclusions to address business questions.

- ***Structured data***: data that adheres to a predefined data model in a tabular format.
- ***Unstructured data***: data that does *not* adhere to a predefined data format.
- ***Big data***: Refers to datasets that are too large and complex for businesses' existing systems to handle utilizing their traditional capabilities to capture, store, manage, and analyze these datasets.

1.2. Why does data analytics matter?

PwC's Annual Global CEO Survey, 86 percent of chief executive officers (CEOs) say they find it important to:

- Champion digital technologies and
- emphasize a clear vision of using technology for a competitive advantage.

85 percent say they put a high value on data analytics.

🔗 PwC's Global Annual CEO Survey report archive

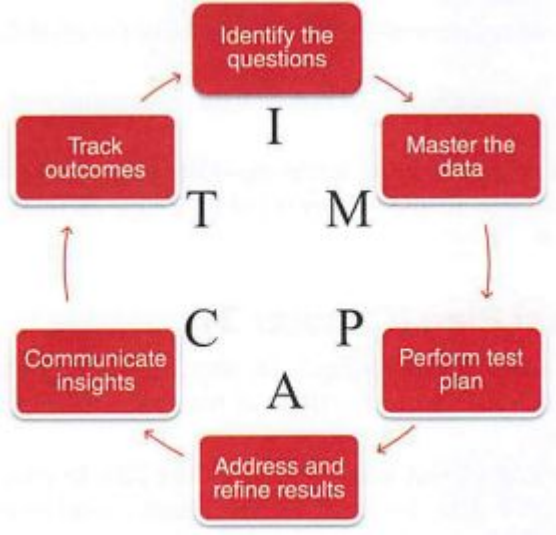
1.3. Why does data analytics matter to accountants?

Audit
In a recent Forbes Insights/KPMG report, "the vast majority of survey respondents believe both that: (a) Audits must better embrace technology, and (b) technology will enhance the quality, transparency, and accuracy of the audit.

Management accounting
Data Analytics are most akin to management accounting. Management accountants (a) are asked questions by management, (b) find data to address those questions, (c) analyze the data, and (d) report the results to management to aid in their decision making.

Financial reporting, financial statement, and tax analyses
Analytics can be used improve the quality of accounting estimates and valuations, assess receivables for collectibility, assess inventory for obsolescence, assess goodwill for impairment, and to address tax planning issues, among other uses.

1.4. The IMPACT cycle



- ***Identify the question(s)***: Are employees circumventing internal controls over payments?
- ***Master the data***: Review data availability in a firm's internal systems.
- ***Perform the test plan***: Implement descriptive, diagnostic, predictive, and/or prescriptive analytics techniques.
- ***Address and refine results***: slice, dice, and manipulate the data; find correlations; test hypotheses; ask ourselves further, hopefully better questions; ask colleagues what they think; and revise and rerun the analysis potentially multiple times.
- ***Communicate insights and Track outcomes***

1.5. Skills required for today's accountants

Accountants don't need to become data scientists, but they must know how to do the following:

- Clearly ***articulate the business problem*** the company is facing.
- ***Communicate*** with the data scientists about specific data needs and understand the underlying quality of the data.
- ***Draw appropriate conclusions*** to the business problem based on the data and make recommendations on a timely basis.
- ***Present their results*** to individual members of management in an accessible manner to each member.

1.5. Skills required for today's accountants

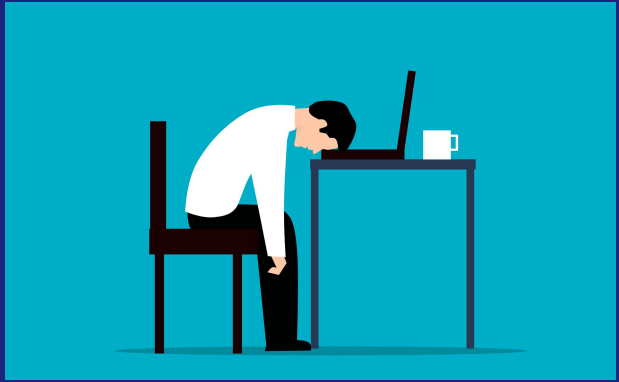
Accountants don't need to become data scientists, but they must know how to do the following:

- Clearly ***articulate the business problem*** the company is facing.
- ***Communicate*** with the data scientists about specific data needs and understand the underlying quality of the data.
- ***Draw appropriate conclusions*** to the business problem based on the data and make recommendations on a timely basis.
- ***Present their results*** to individual members of management in an accessible manner to each member.

1.5. Skills required for today's accountants

Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms

Hands on example: Journal entries to trial balance



Data file

2. Mastering the data

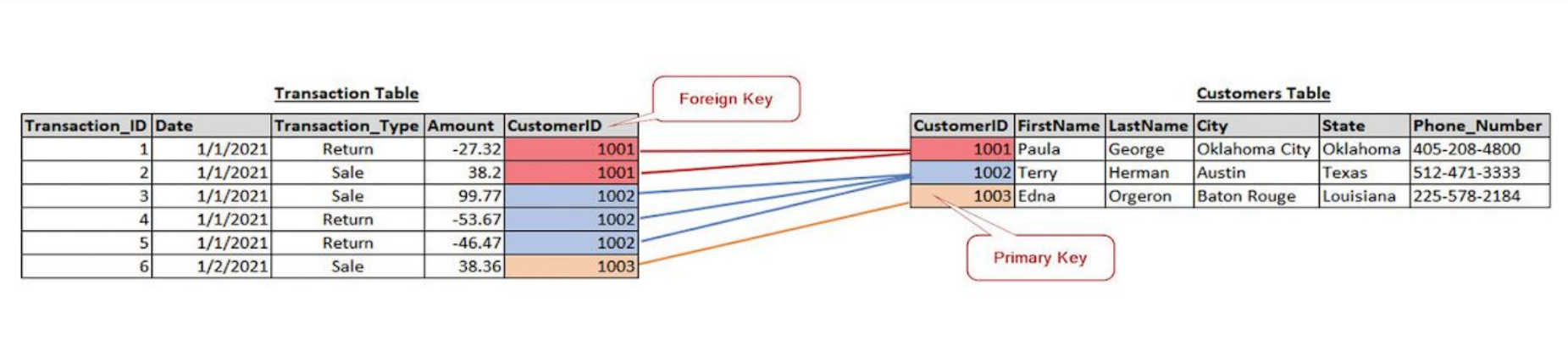
2.1. Internal and external data sources

Internal sources	Description
Enterprise resource planning (ERP)	A category of business management software that integrates applications from throughout the business (such as manufacturing, accounting, finance, human resources, etc.) into one system.
Accounting information system (AIS)	A system that records, processes, reports, and communicates the results of business transactions to provide financial and nonfinancial information for decision-making purposes.
Supply chain management (SCM) system	Includes information on active vendors (their contact info, where payment should be made, how much should be paid), the orders made to date (how much, when the orders are made) or demand schedules for what component of the final product is needed when.
Customer relationship management (CRM) system	An information system for overseeing all interactions with current and potential customers with the goal of improving relationships.
Human resource management (HRM) system	An information system for managing all interactions with current and potential employees.

2.1. Internal and external data sources

Category	Dataset Description	Website
Economics	BRICS World Bank Indicators (Brazil, Russia, India, China and South Africa)	https://www.kaggle.com/docstein/brics-world-bank-indicators
Economics	Bureau of Economic Analysis data	http://www.bls.gov/data/
Financial	Financial statement data	https://www.calcbench.com/
Financial	Financial statement data, EDGAR, Securities and Exchange Commission	https://www.sec.gov/edgar.shtml
Financial	Analyst forecasts	Yahoo! Finance (finance.yahoo.com), Analysis Tab
Financial	Stock market dataset	https://www.kaggle.com/borismarjanovic/price-volume-data-for-all-us-stocks-etfs
Financial	Credit card fraud detection	https://www.kaggle.com/mlg-ulb/creditcardfraud
Financial	Daily News/Stock Market Prediction	https://www.kaggle.com/aaron7sun/stocknews
Financial	Retail Data Analytics	https://www.kaggle.com/manjeetsingh/retaildataset
Financial	Peer-to-peer lending data of approved and rejected loans	lendingclub.com (requires login)
Financial	Daily stock prices (and weekly and monthly)	Yahoo! Finance (finance.yahoo.com), Historical Data Tab
Financial	Financial and economic summaries by industry	http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datacurrent.html
General	data.world	https://data.world/
General	kaggle.com	https://www.kaggle.com/datasets
Government	State of Ohio financial data (Data Ohio)	https://data.ohio.gov/wps/portal/gov/data/
Government	City of Chicago financial data	https://data.cityofchicago.org
Government	City of New York financial data	http://www.checkbooknyc.com/spending_landing/yeartype/B/year/119
Marketing	Amazon product reviews	https://data.world/datafiniti/consumer-reviews-of-amazon-products
Other	Restaurant safety	https://data.cityofnewyork.us/Health/DOHMH-New-York-City-Restaurant-Inspection-Results/43nn-pn8j
Other	Citywide payroll data	https://data.cityofnewyork.us/City-Government/Citywide-Payroll-Data-Fiscal-Year-/k397-673e
Other	Property valuation/assessment	https://data.cityofnewyork.us/City-Government/Property-Valuation-and-Assessment-Data/yjxr-fw8i
Other	USA facts—our country in numbers	https://www.irs.gov/uac/tax-stats
Other	Interesting fun datasets—14 data science projects with data	https://towardsdatascience.com/14-data-science-projects-to-do-during-your-14-day-quarantine-8bd60d1e55e1
Other	Links to Big Data Sets—Amazon Web Services	https://aws.amazon.com/public-datasets/
Real Estate	New York Airbnb data explanation	https://www.kaggle.com/dgomonov/new-york-city-airbnb-open-data
Real Estate	U.S. Airbnb data	https://www.kaggle.com/kritikseth/us-airbnb-open-data/tasks?taskId=2542
Real Estate	TripAdvisor hotel reviews	https://www.kaggle.com/andrewmvd/trip-advisor-hotel-reviews
Retail	Retail sales forecasting	https://www.kaggle.com/tevecsystems/retail-sales-forecasting

2.2. Relational databases



2.2. Relational databases

Relational databases break data into separate tables, each containing a unique list of the items stored. A relational database is comprised of:

- ***Tables***: Data organized into sets of columns (fields) and rows (records).
- ***Fields (aka Variables)***: The columns that contain descriptive characteristics about the observations in the table.
- ***Records***: The rows, with each observation corresponding to a record, or unique instance, of what is being described in the table.

A ***primary key*** is a field (variable) in a relational database that serves as a unique identifier.

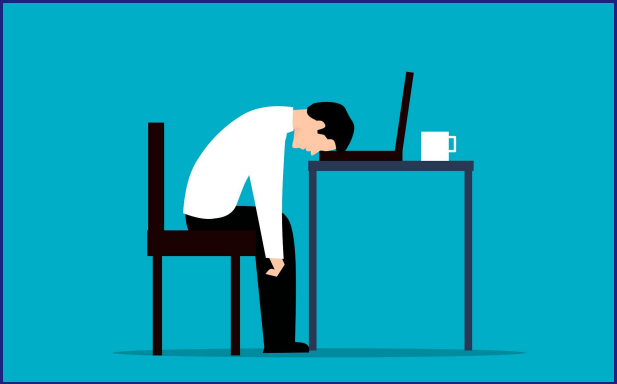
2.3. The Extract, Transform, Load (ETL) process

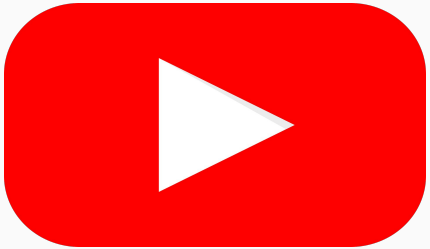
	Description
Extract	Step 1: Determine the purpose and scope of the data request Step 2: Obtain the data
Transform	Step 3: Validating the data for completeness and integrity Step 4: Cleaning the data
Load	Step 5: Loading the data for data analysis

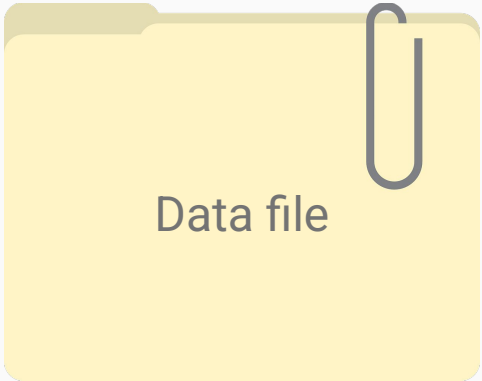
2.4. Ethical considerations

- How does the company use data, and to what extent are they integrated into firm strategy?
- Does the company send a privacy notice to individuals when their personal data are collected?
- Does the company assess the risks linked to the specific type of data the company uses?
- Does the company have safeguards in place to mitigate the risks of data misuse?
- Does the company have the appropriate tools to manage the risks of data misuse?
- Does our company conduct appropriate due diligence when sharing with or acquiring data from third parties?

Hands on example: Inventory management by customer

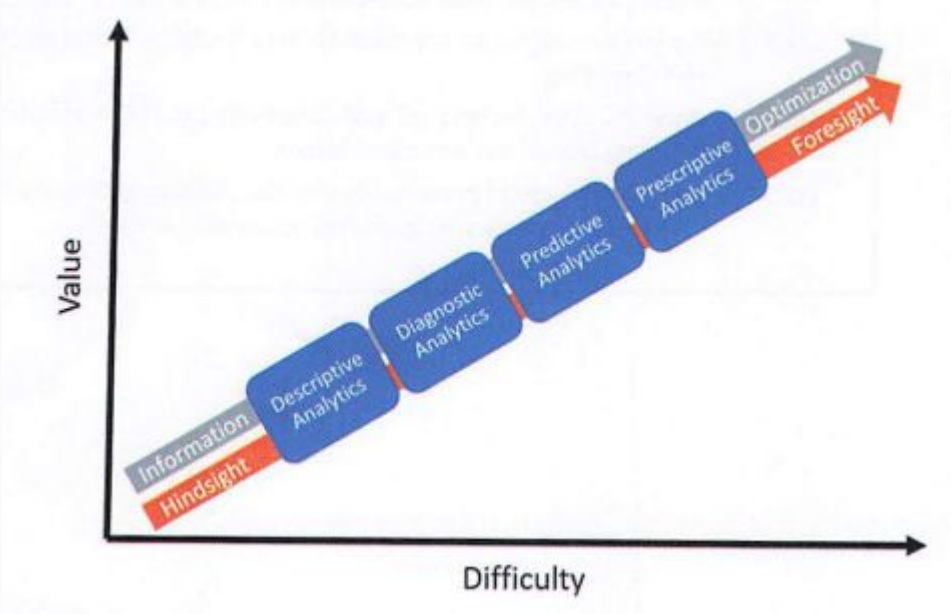






3. Performing the test plan and analyzing results

3.1. Four types of data analytics



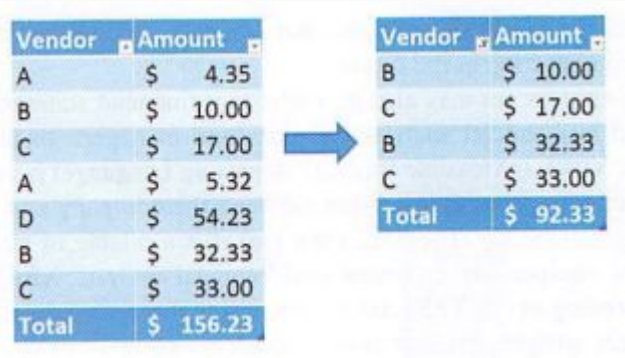
3.1. Four types of data analytics

Descriptive	Procedures that summarize existing data to determine what has happened in the past.
Diagnostic	Procedures that explore the current data to determine why something has happened the way it has, typically comparing the data to a benchmark.
Predictive	Procedures used to generate a model that can be used to determine what is likely to happen in the future.
Prescriptive	Procedures that work to identify the best possible options given constraints or changing conditions.

3.2. Descriptive analytics

Statistic	Excel formula	Description
Sum	=SUM()	The total value of all numerical values
Mean	=AVERAGE()	The center value; sum of all observations divided by the number of observations
Median	=MEDIAN()	The middle value that divides the top half of the data from the bottom half
Minimum	=MIN()	The smallest value
Maximum	=MAX()	The largest value
Count	=COUNT()	The number of observations
Frequency	=FREQUENCY()	The number of observations in each of a series of numerical or categorical buckets
Standard deviation	=STDEV()	The variability or spread of the data from the mean; a larger standard deviation means a wider spread away from the mean
Quartile	=QUARTILE()	The value that divides a quarter of the data from the rest; indicates skewness of the data
Correlation coefficient	=CORREL()	How closely two datasets are correlated or predictive of each other

3.2. Descriptive analytics



Vendor	Amount
A	$ 4.35
B	$ 10.00
C	$ 17.00
A	$ 5.32
D	$ 54.23
B	$ 32.33
C	$ 33.00
Total	$ 156.23

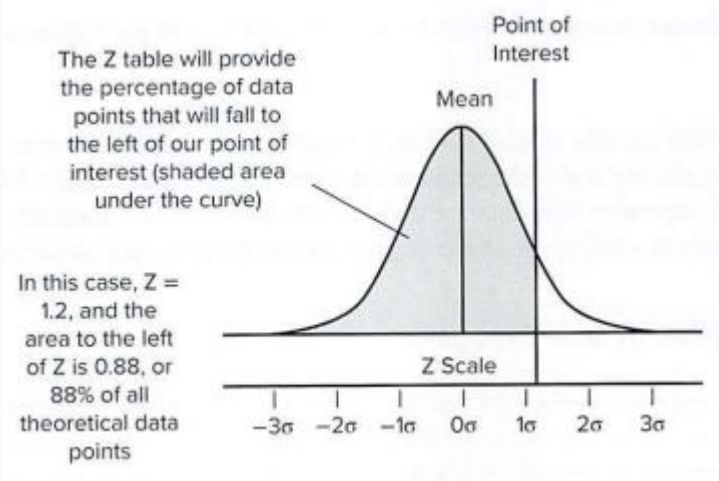
Vendor	Amount
B	$ 10.00
C	$ 17.00
B	$ 32.33
C	$ 33.00
Total	$ 92.33

Data reduction attempts to reduce the amount of detailed information considered to focus on the most critical, interesting, or abnormal items (e.g., highest cost, highest risk, largest impact, etc.). It does this by filtering through a large set of data (perhaps the total population) and reducing it to a smaller set that has the vast majority of the critical information of the larger set. The data reduction approach is done primarily using structured data - that is, data that are stored in a database or spreadsheet and are readily searchable.

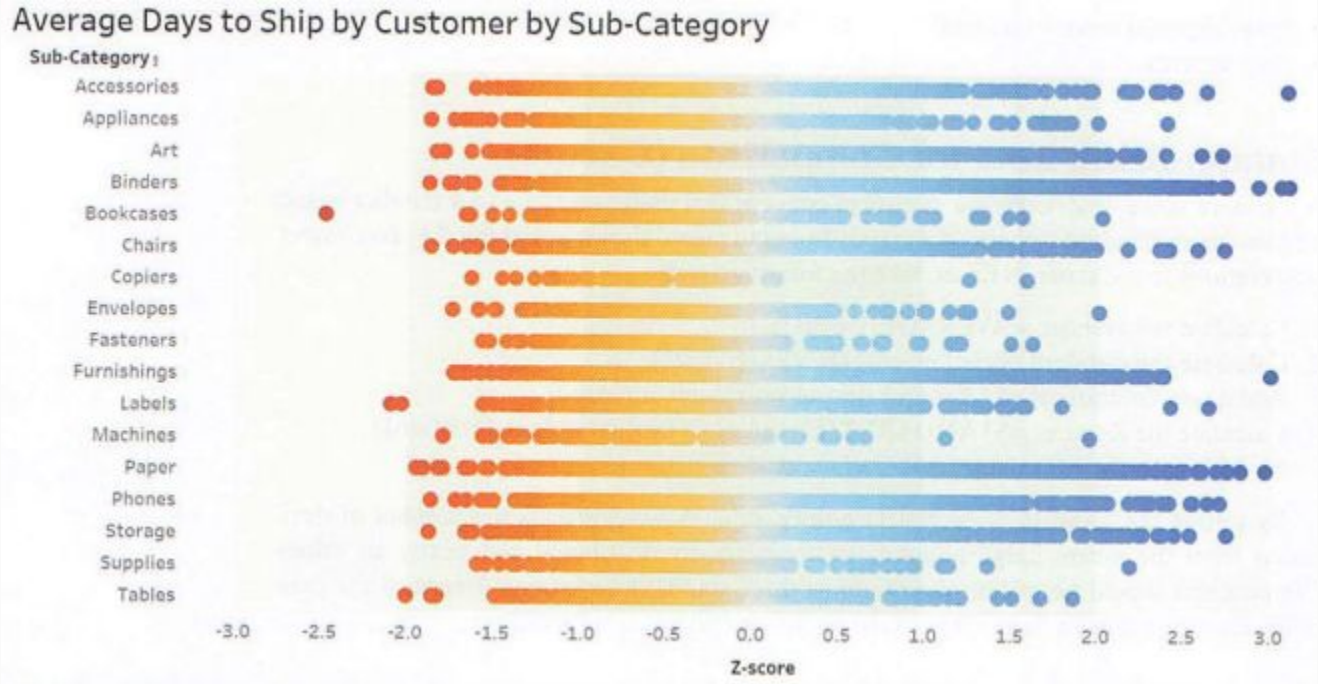
3.3. Diagnostic analytics

Standardizing Data for Comparison (Z-score)

A standard score or Z-score is a statistical concept that assigns a value to a number based on how many standard deviations it stands from the mean.

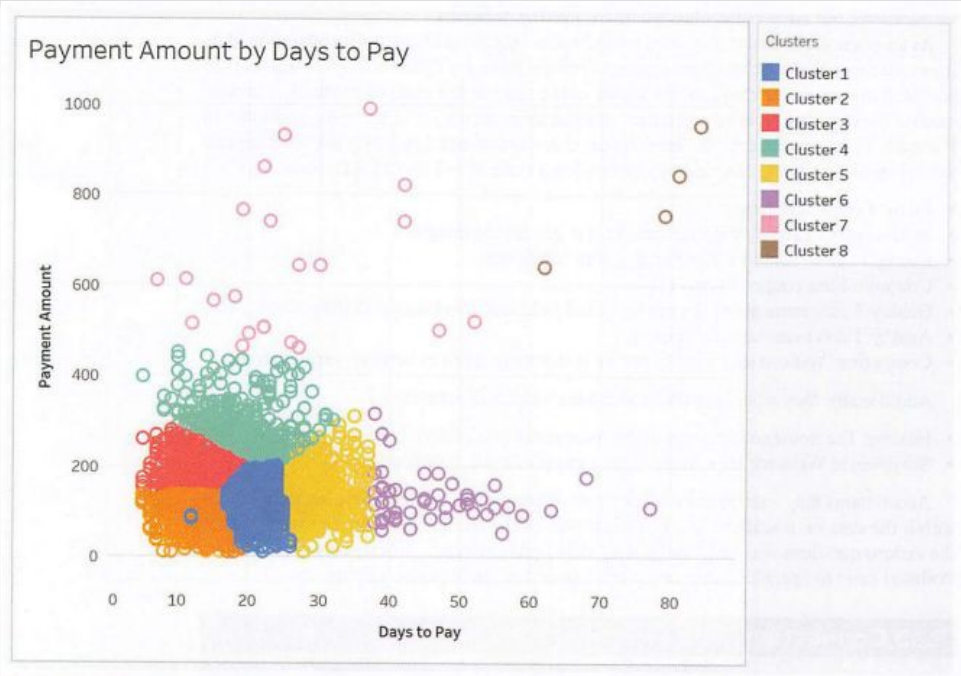


3.3. Diagnostic analytics

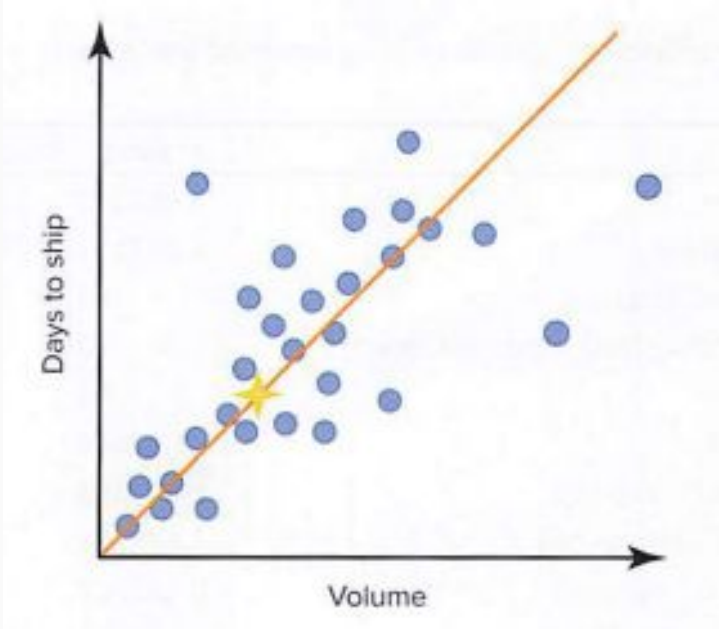


3.3. Diagnostic analytics

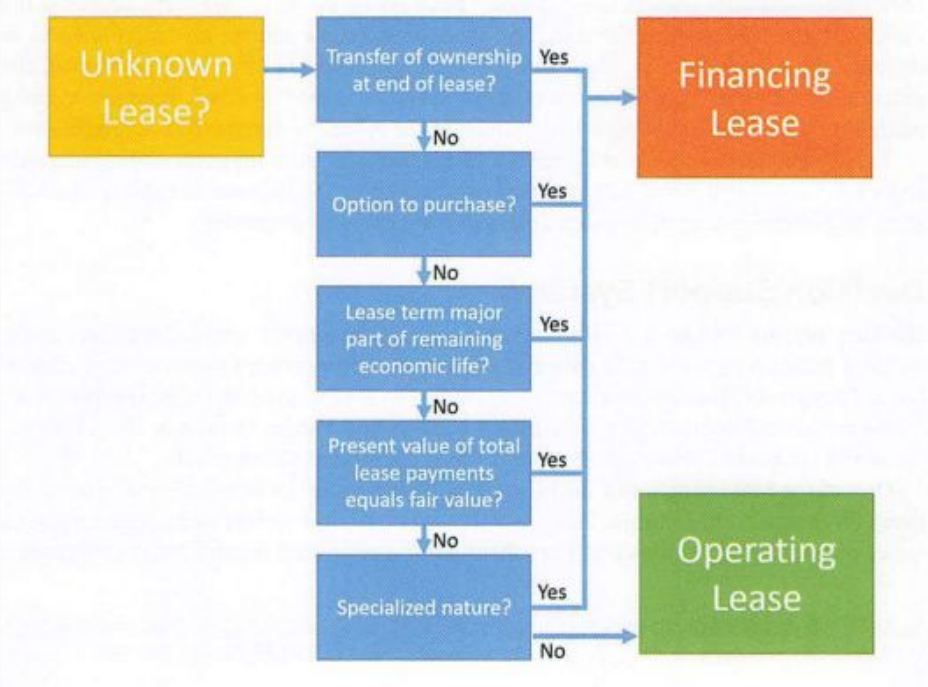
Clustering
The clustering data approach works to identify groups of similar data elements and the underlying relationships of those groups.



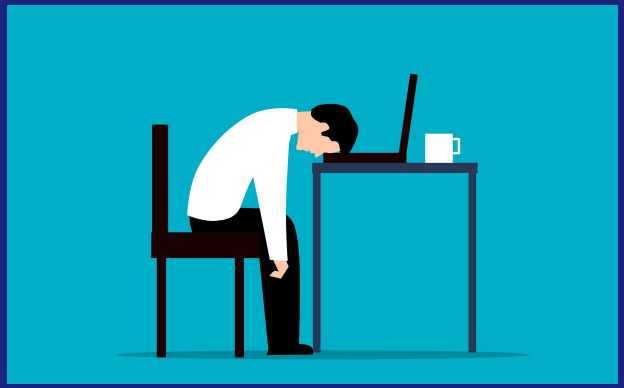
3.4. Predictive analytics

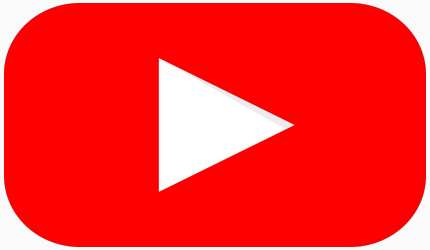


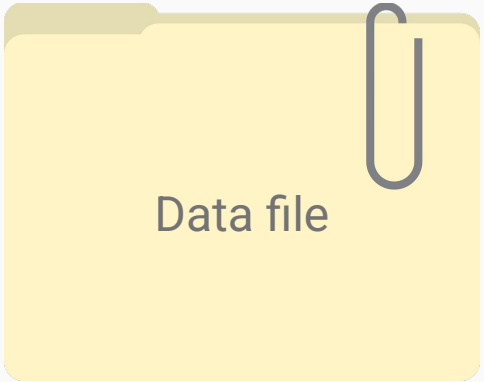
3.5. Prescriptive analytics



Hands on example: Predicting Bankruptcy Using Altman's Z Score





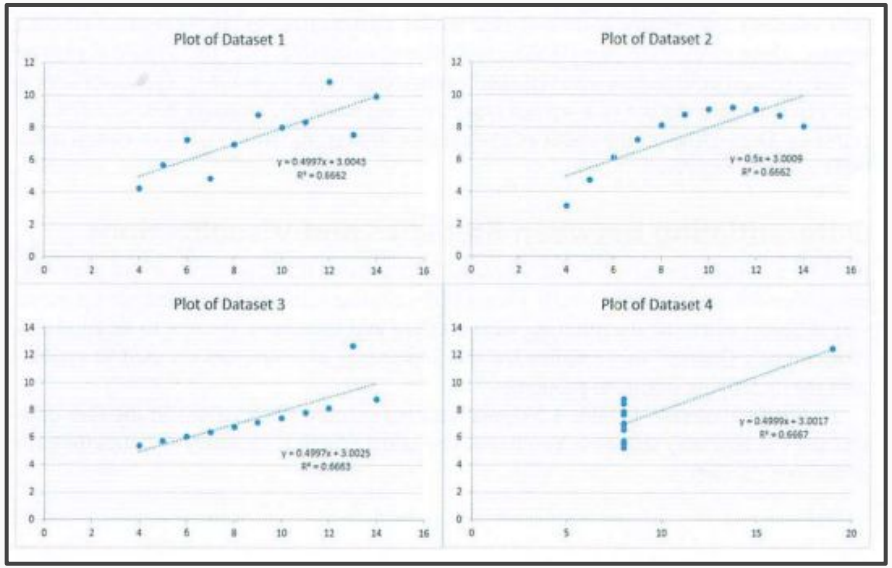


4. Communicating results and visualizations

4.1. Communicating results

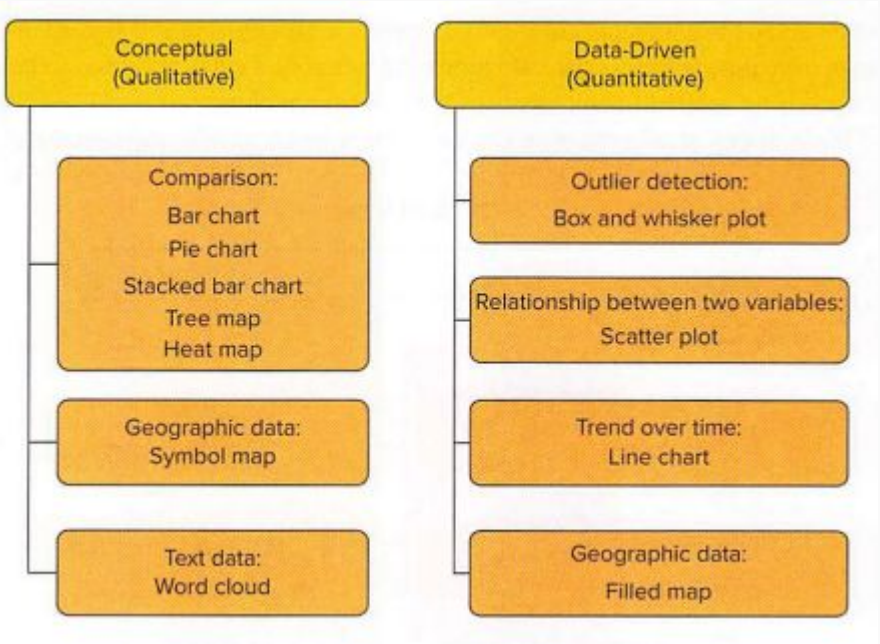
Anscombe's Quartet (Data)

Observation Number	Dataset 1 x1	Dataset 1 y1	Dataset 2 x2	Dataset 2 y2	Dataset 3 x3	Dataset 3 y3	Dataset 4 x4	Dataset 4 y4
1	10	8.04	10	9.14	10	7.46	8	6.58
2	8	6.95	8	8.14	8	6.77	8	5.76
3	13	7.58	13	8.74	13	12.74	8	7.71
4	9	8.81	9	8.77	9	7.11	8	8.84
5	11	8.33	11	9.26	11	7.81	8	8.47
6	14	9.96	14	8.1	14	8.84	8	7.04
7	6	7.24	6	6.13	6	6.08	8	5.25
8	4	4.26	4	3.1	4	5.39	19	12.5
9	12	10.84	12	9.13	12	8.15	8	5.56
10	7	4.82	7	7.26	7	6.42	8	7.91
11	5	5.69	5	4.74	5	5.73	8	6.89
Summary Statistics								
N (count	11	11	11	11	11	11	11	11
mean (average)	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
SD (Standard Deviation)	3.32	2.03	3.32	2.03	3.32	2.03	3.32	2.03
r (Correlation)	0.82		0.82		0.82		0.82	

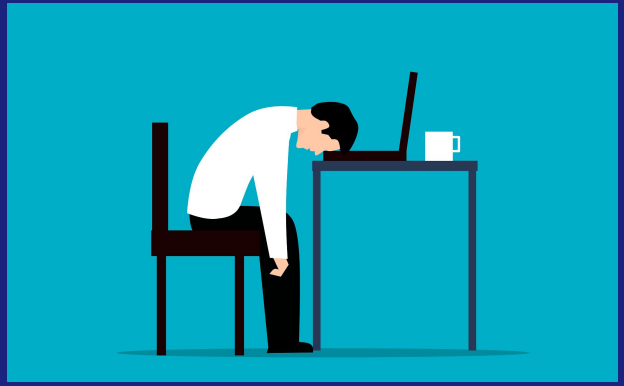


4.2. Determine the purpose of your data visualization

4.3. Choosing the best charts



Hands on example: Create a Dashboard in Excel Using PivotTables and Slicers



Data file

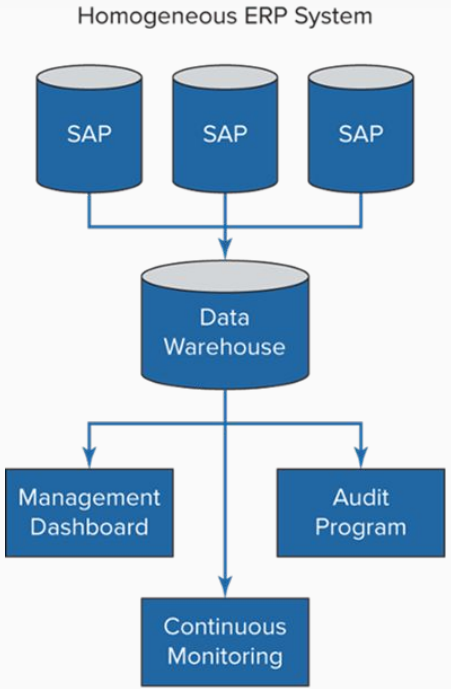
5. The modern accounting environment

5.1. Automation and its impact on decision making

Businesses have embraced automation to capture transactions, metadata (for example, timestamps, user details, and unstructured data).

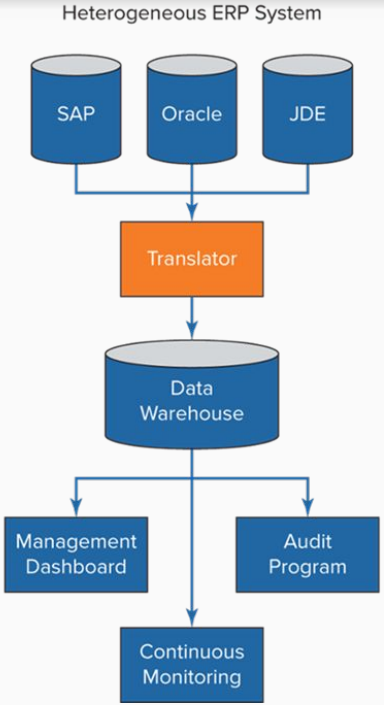
5.2. Enterprise data and data models

Firm data stored in ***homogeneous ERP systems*** is easier to work with because the location of key data is known.



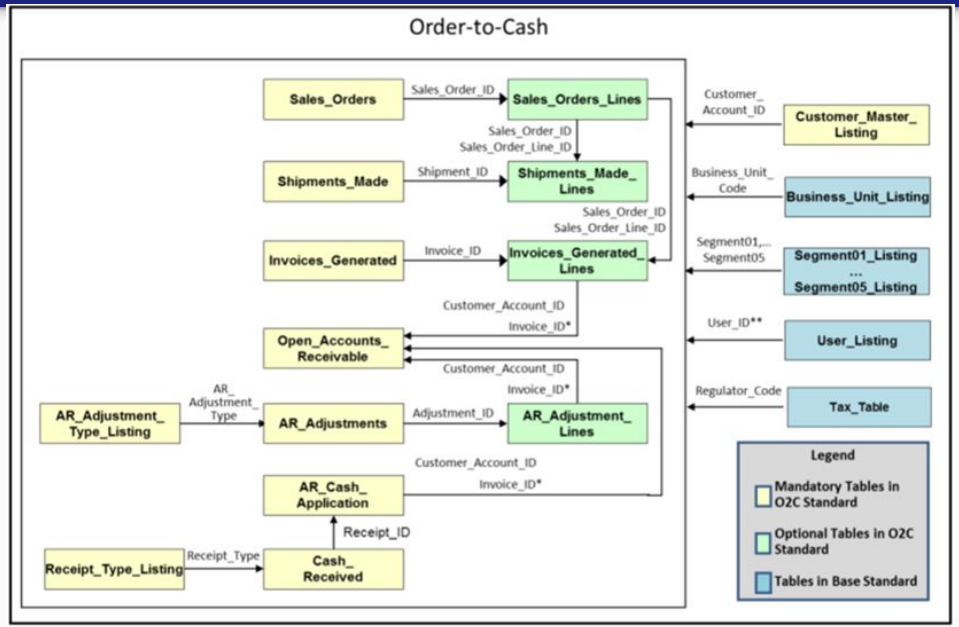
5.2. Enterprise data and data models

Firm data in ***heterogeneous ERP systems*** requires a systems translation engine and mapping and is usually the result of acquisitions.



5.2. Enterprise data and data models

The AICPA's ***Audit Data Standards*** identify common tables and fields needed for audit analytics.



5.3. Approaches to automating procedures

Most of the effort in data analytics is used to identify data, map tables and fields, and develop visualizations.

Once identified and developed, automation involves identifying the timing of updates, parameters, and what to do when an outlier is detected.

5.4. Continuous monitoring techniques

Data Analytics and audit automation allows auditors to continuously monitor systems and processes.

Continuous auditing provides real-time assurance over business processes. This provides increased coverage and timeliness of audit procedures.

Continuous monitoring evaluates controls and transactions and is primarily used by management to show effectiveness of internal controls.

Continuous reporting provides the status of the audit procedures and output of the information systems.

6. Audit data analytics

6.1. Different types of analysis for auditing

Data Analytics can be applied to the auditing function to increase coverage of the audit, while reducing the time the auditor dedicates to the audit tasks.

6.2. Descriptive analytic techniques in auditing

Descriptive analytics are useful for sorting and summarizing data to create a baseline or point of reference for more advanced analytics.

This substantive test of account balances evaluates the date of an order and groups it into buckets.

Extremely old accounts that haven't been resolved should be flagged by the auditor.

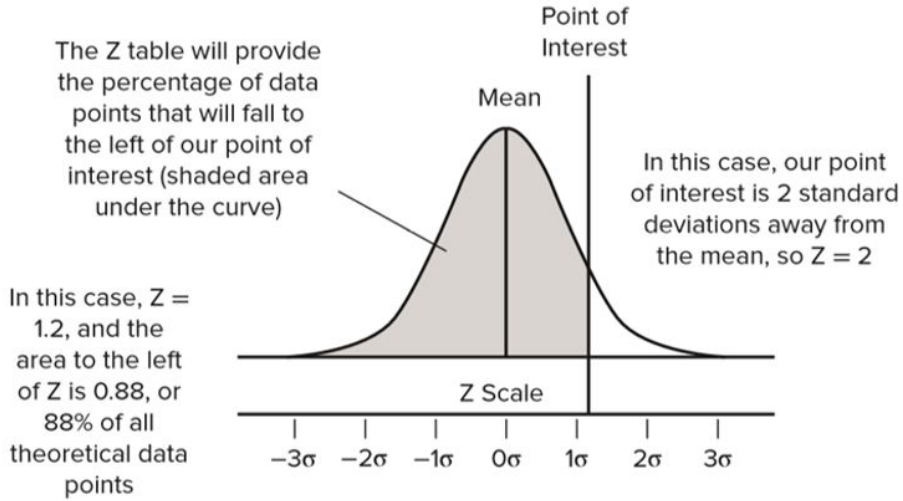
Exhibit 6-3 Aging of Accounts Receivable

Days old	Total
0 to 30	154,322
31 to 60	74,539
61 to 90	42,200
>90	16,900

6.3. Diagnostic techniques used in auditing

- High Z-score values represent outliers.
- A score above 3 standard deviations is very rare.

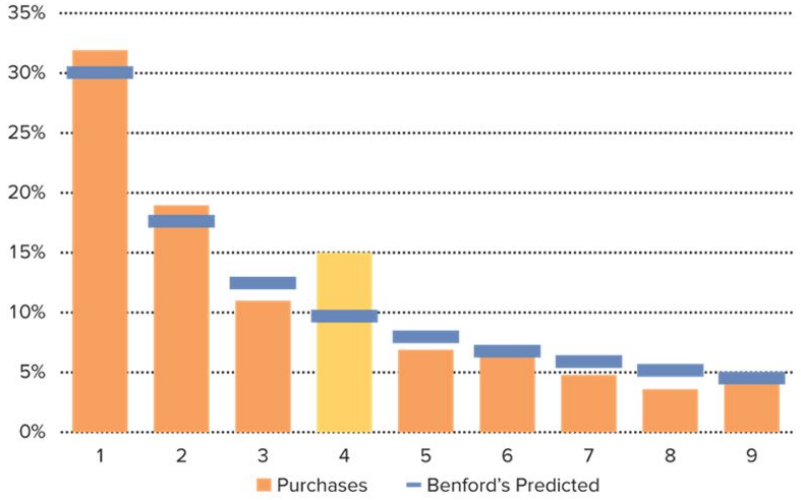
Exhibit 6-4 The Z-score shows the relative position of a point of interest to the population.



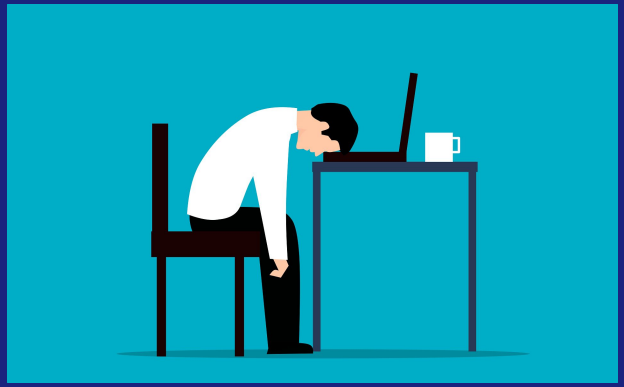
6.3. Diagnostic techniques used in auditing

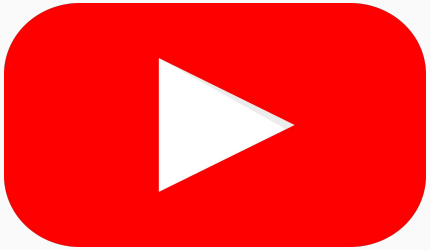
- In large sets of numbers, the first digit follows a predictable distribution.
- It can be used to detect an abnormal volume of transactions that begin with a specific digit.
- Evaluate the average Benford's Law value to identify individuals.

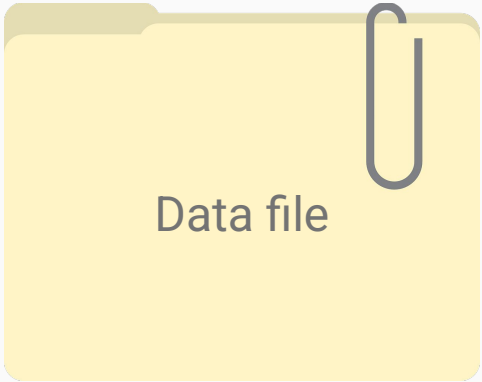
Exhibit 6-6 Structured purchases may look normal, but they alter the distribution under Benford's law.



Hands on example: Fuzzy Matching and Fake Employees / Vendors







7. Managerial analytics

Examples of management accounting questions

- What percentage of the airline company's departures were on time this past month?
- What was the segment margin for the West Coast and Midwest regions last quarter?
- Which products are the most profitable for the company? How much did Job #304 cost?
- Why is segment margin higher on the West Coast than in the Midwest?
- Why did our rate of production defects go down this month compared to last month?
- What is driving the price variance and labor rate variance?

7.2. Descriptive and diagnostic analytics in management accounting

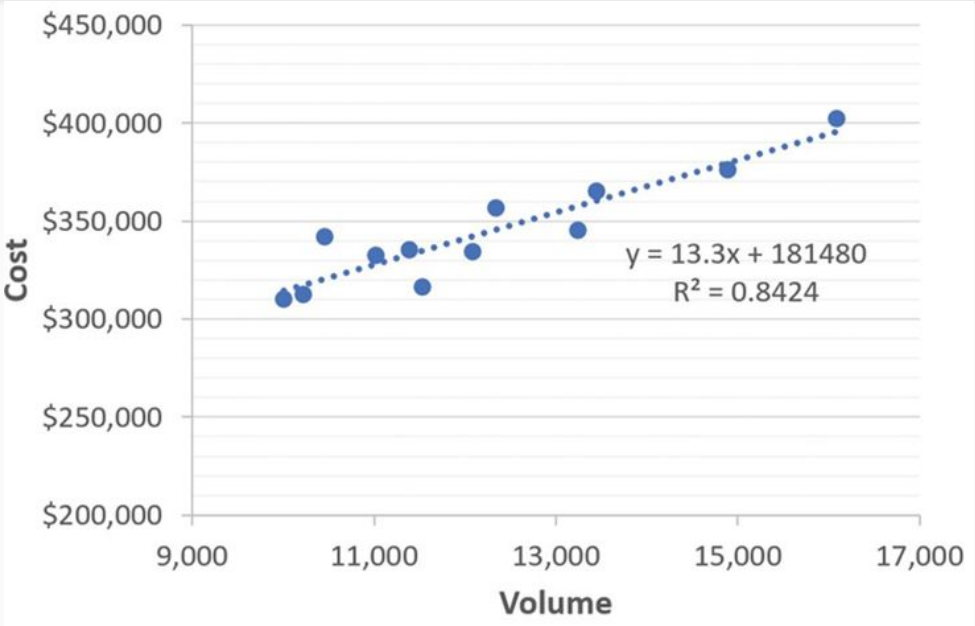
- Managers compare actual results to budgeted results to determine whether a variance is favorable or unfavorable.

- Bullet charts help managers identify root causes of the variance (for example, the price we pay for a raw material or the increased volume of sales) and drill-down to determine the good performance to replicate and the poor performance to eliminate.

7.2. Descriptive and diagnostic analytics in management accounting

- Managers must also understand what is driving the costs and profits to plan for the future and apply to budgets or use as input for lean accounting processes.

- Predictive analytics, such as regression analysis, might evaluate actual production volume and total costs to estimate the mixed cost line equation.



7.3. KPIs as part of a balanced scorecard

Key performance indicators (KPIs) are a specific type of performance metrics used to measure performance at a company.

The **Balanced Scorecard** identifies the most important metrics to measure and target goals for comparison.

7.3. KPIs as part of a balanced scorecard

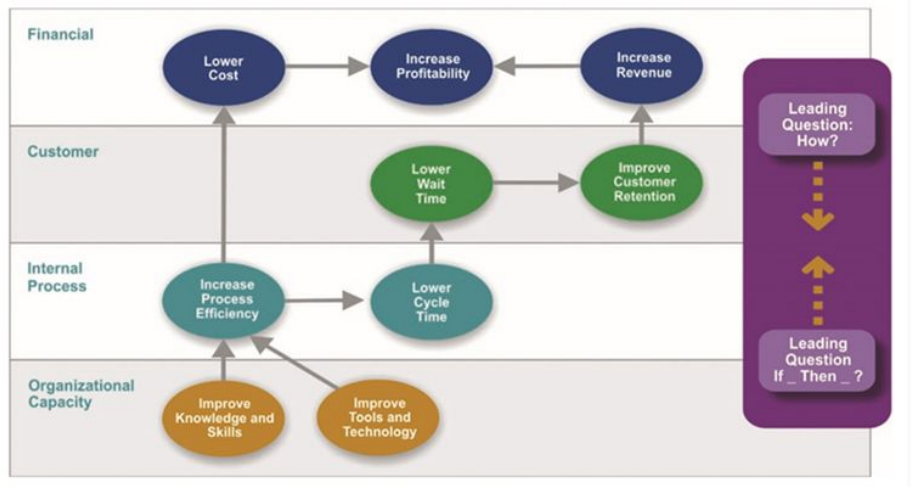
Four components of balanced scorecard

Financial: how the company generates value.

Customer: how the company interacts with customers.

Internal process: how efficiently the company is operating.

Organizational capacity: how the company is training employees.



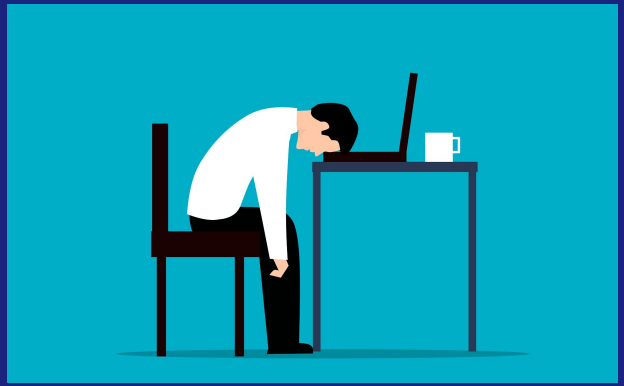
7.5. Address and refine results

The information created from your analysis will need to be refined for the use case.

Consider these questions:

- Which metric are you using most frequently to help you make decisions?
- Are there any metrics that you do not use? If so, why aren't they helpful?
- Are you downloading the data to do any additional analysis after working with the dashboard, and if so, can the dashboard be improved to save those extra steps?
- Are there any metrics that should be available on the dashboard to help you with decision making?

Hands on example: Evaluating the Relationship between Sales and Advertising Expense



Data file

8. Financial statement analytics

8.1. Types of financial statement analysis

Vertical Analysis of a Common Size Financial Statement

Apple Inc

	2017	2018	$ Change
Revenue	$ 96,571	$ 110,360	$ 13,789
Cost of revenue	$ 34,261	$ 38,353	$ 4,092
Gross profit	$ 62,310	$ 72,007	$ 9,697
Total operating expenses	$ 32,979	$ 36,949	$ 3,970
Operating income	$ 29,025	$ 35,058	$ 6,033
Other expenses	$ 3,536	$ 18,487	$ 14,951
Net income	$ 25,489	$ 16,571	$ (8,918)

Microsoft Corp

	2019	2020	Vertical
Revenue	$ 229,234	$ 265,595	100.00%
Cost of revenue	$ 141,048	$ 163,756	61.66%
Gross profit	$ 88,186	$ 101,839	38.34%
Total operating expenses	$ 26,842	$ 30,941	11.65%
Operating income	$ 61,344	$ 70,898	26.69%
Other expenses	$ 12,993	$ 11,367	4.28%
Net income	$ 48,351	$ 59,531	22.41%

8.1. Types of financial statement analysis

Comparison of Ratios among Microsoft (MSFT), Apple (AAPL), and Facebook (FB)

Microsoft Corp

Trend	2018	2019	2020		
	0.19	0.42	0.40	$ 44,281	$ 110,317
	0.15	0.31	0.31	$ 44,281	$ 143,015
	0.43	0.46	0.49	$ 143,015	$ 293,934
	2.99	2.95	2.66	$ 293,934	$ 110,317

Return on equity (ROE) = Net income / Average shareholders' equity

Profit margin (PM) = Net income / Net sales

Asset turnover (AT) = Net sales / Average total assets

Equity multiplier (EM) = Average total assets / Average total equity

DuPont Framework

ROE = PM x AT x EM

Market Comparison

	MSFT	AAPL	FB	Comp.
Return on equity (ROE)	0.40	0.74	0.25	
Profit margin (PM)	0.31	0.21	1.75	
Asset turnover (AT)	0.49	0.83	0.11	
Equity multiplier (EM)	2.66	4.25	1.28	
DuPont Framework	0.40	0.74	0.25	

8.2. Create visualizations

Microsoft Corp	2016	2017	2018	2019	2020	
Revenue	$ 91,154	$ 96,571	$ 110,360	$ 125,843	$ 143,015	
Cost of revenue	$ 32,780	$ 34,261	$ 38,353	$ 42,910	$ 46,078	
Gross profit	$ 58,374	$ 62,310	$ 72,007	$ 82,933	$ 96,937	
Total operating expenses	$ 31,186	$ 33,285	$ 36,949	$ 39,974	$ 43,978	
Operating income	$ 26,078	$ 29,025	$ 35,058	$ 42,959	$ 52,959	
Other expenses	$ 5,539	$ 3,536	$ 18,487	$ 3,719	$ 8,678	
Net income	$ 20,539	$ 25,489	$ 16,571	$ 39,240	$ 44,281	

8.2. Create visualizations

8.3. Text mining and sentiment analysis

Text mining analyzes the frequency of words in unstructured data (for example, financial disclosure) and matches those to a sentiment dictionary (for example, words identified as positive or negative).

8.3. Text mining and sentiment analysis

Positive

- (e.g., enable, achieve)

Negative

- (e.g., loss, adverse)

Uncertain

- (e.g., anticipate, depend)

Litigious

- (e.g., mediate, petition)

Modal

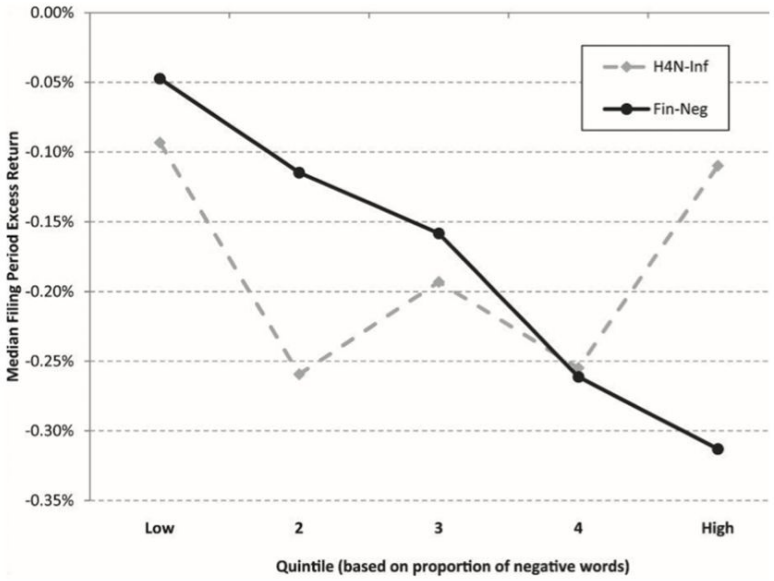
- (e.g., possible, likely)

Constraining

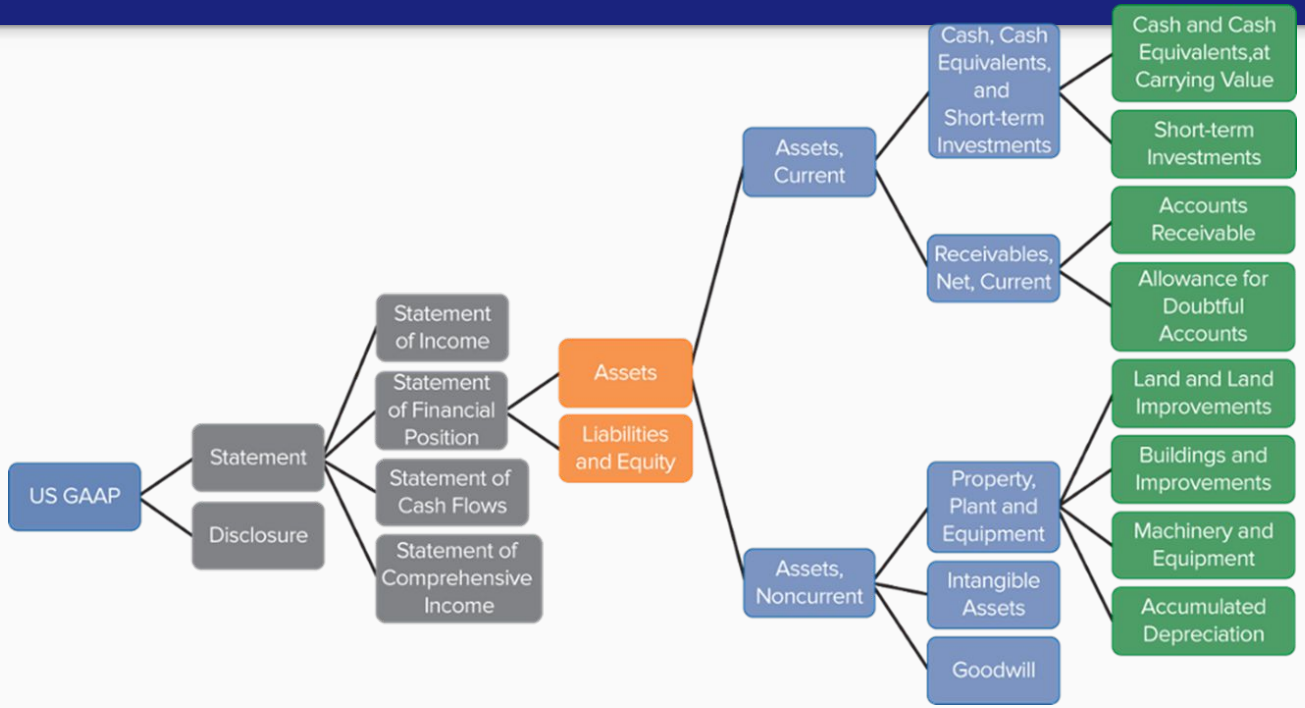
- (e.g., commit, impair)

8.3. Text mining and sentiment analysis

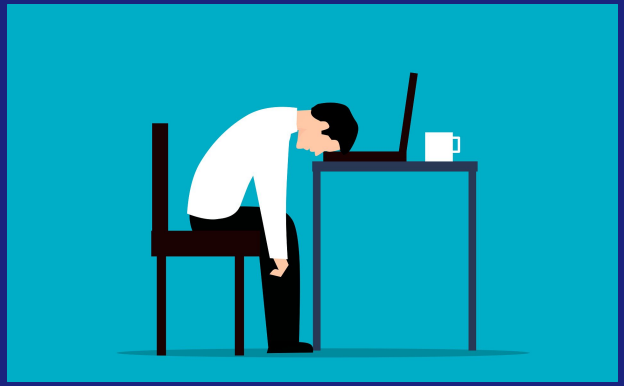
Stock Market Reaction (Excess Return) of Companies Sorted by Proportion of Negative Words



8.4. XBRL tagging



Hands on example: DuPont Analysis of Financial Performance



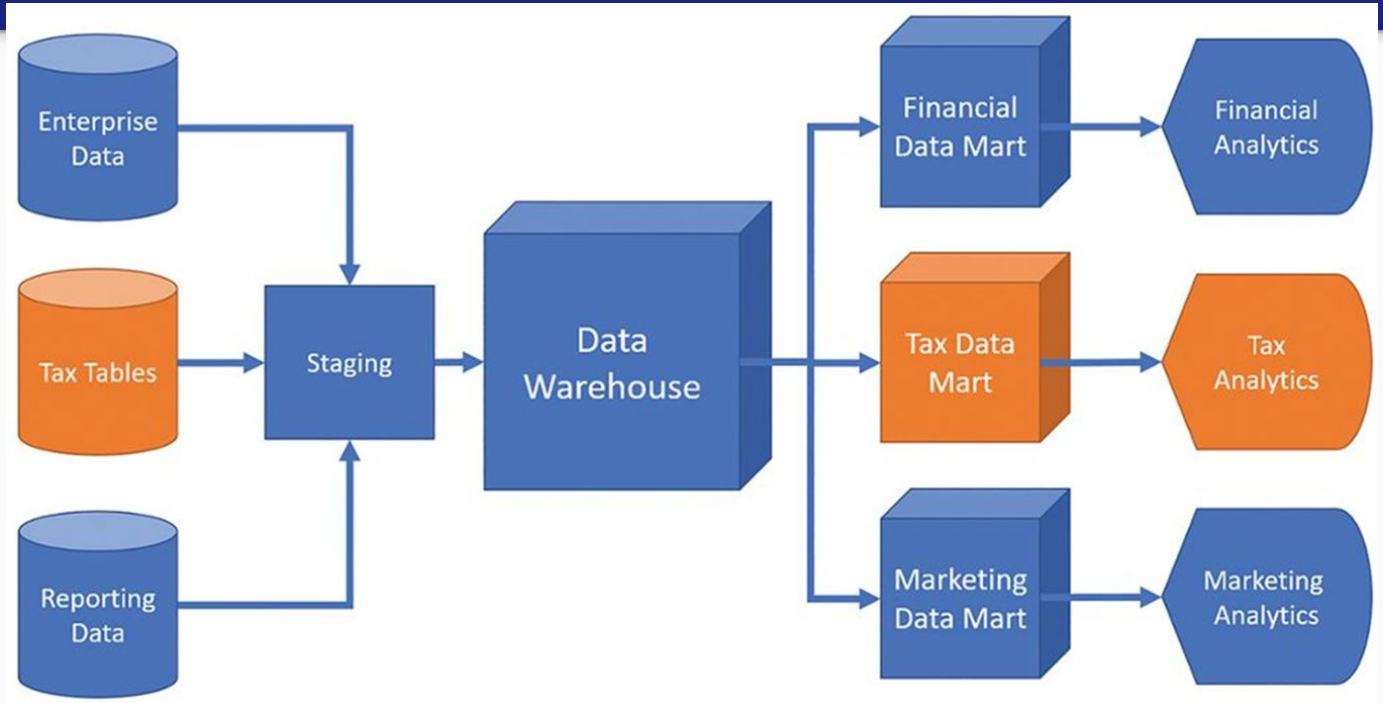
Data file

9. Tax analytics

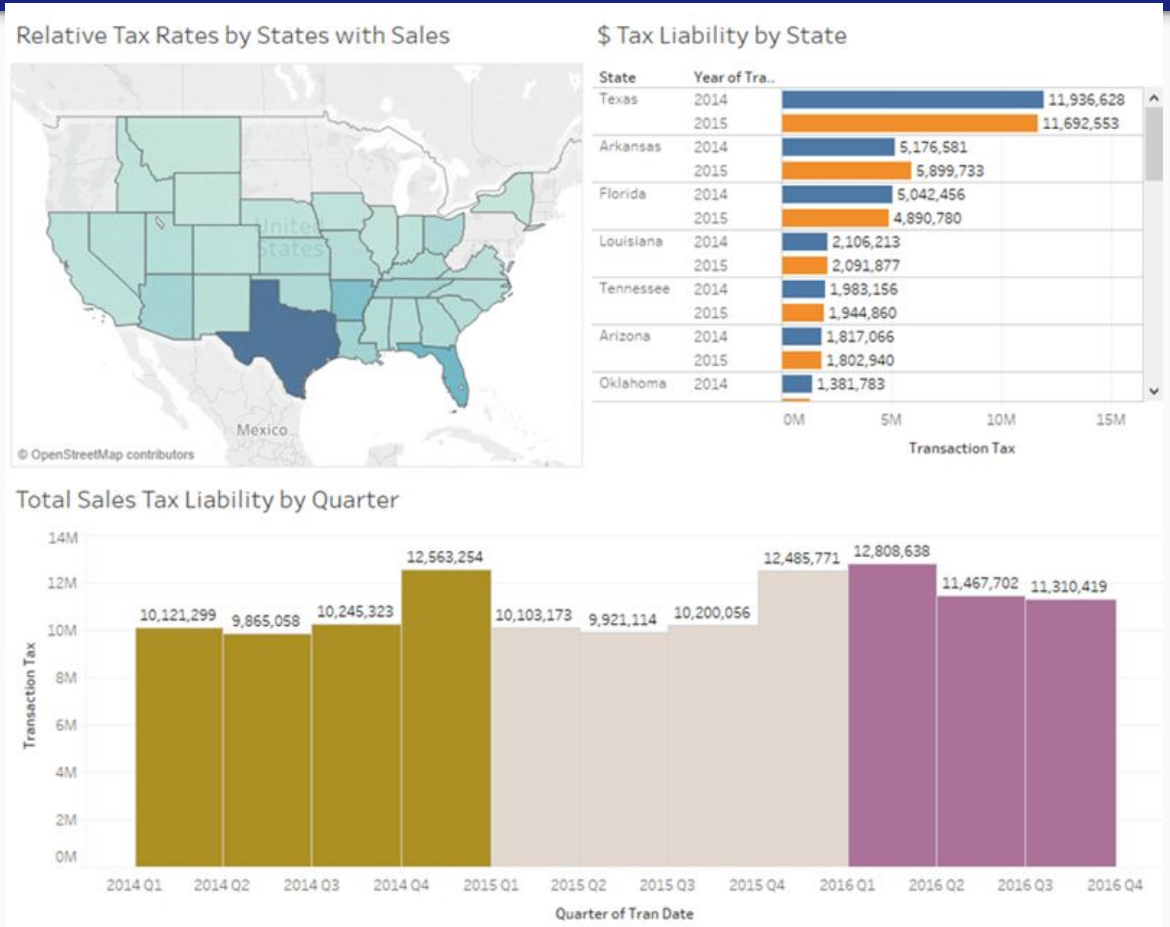
9.1. Types of tax analytics problems

- What is the amount of tax paid each year by entity (nationwide, corporate, and individual) or tax category (income, sales, property, excise, etc.)?
- What is the difference between GAAP-basis and taxable income (book-tax differences)?
- What is the amount of sales tax paid compared to expectations?
- What is the amount of R&D tax credit we expect to qualify for in the future?
- If certain tax legislation passes, what level of exposure (additional tax) might the company face?
- What will be the amount of taxes we owe if we pursue a merger or acquisition?

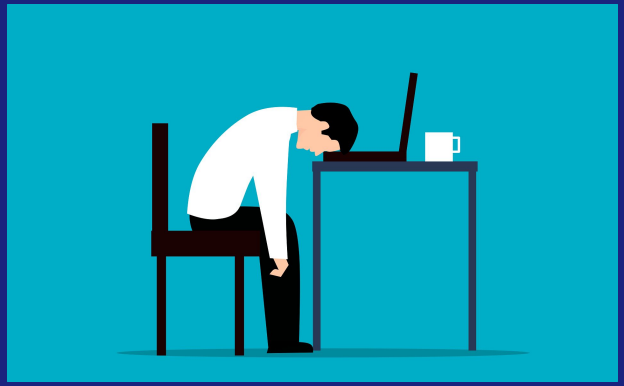
9.2. Tax data sources



9.3. Visualizations in tax analytics



Hands on example: Linking Two Tables Using VLOOKUP for State Tax Rates



Data file