Appendix B
Data Analytics

Road Map

Learning Objectives		Page	eLecture	Assignments
1	Identify and define the four types of data analytics.	519	eF.1	1, 6, 7, 8,
2	Describe the use of data analytics within the accounting profession.	520	eF.2	1, 6, 7, 8, 13, 18,
3	Describe the analytics mindset.	521	eF.3	2, 3, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25
4	Describe data visualization best practices.	523	eF.4	4, 5, 10, 11, 12, 14, 15, 16, 17, 19, 20, 21, 22, 24

Data Analytics

Data analytics can broadly be defined as the process of examining sets of data with the goal of discovering useful information from patterns found in the data. Increasingly, this process is aided by computers running programs ranging from basic spreadsheet software, such as **Microsoft Excel** and **Google Sheets**, to specialized software, such as **Tableau** or **Power BI**. This technology can reveal trends and insights that would otherwise be lost in the overwhelming amount of data.

Big Data

The concept of data analytics is intertwined with the concept of **big data**. While no precise definition exists for big data, a commonly accepted definition is that big data is a collection of data that is both extremely large and also extremely complex, thus making its analysis beyond the scope of traditional tools. Important attributes of big data, commonly referred to as the four V's, are Volume, Variety, Velocity, and Veracity. **Volume** refers to the amount of data. According to IDC (a market intelligence company), there were 33 available zettabytes of data globally in 2018. IDC predicted that the amount of data would increase to 175 zettabytes by 2025. (Just so you know, there are 21 zeros in one zettabyte.) Total amounts of data are growing because we are creating more data (through new technologies) and because we are able to store more data (using cloud storage services like Amazon Web services [AWS] and Microsoft Azure). Massive data sets can't be managed on a single machine. They must be stored in clusters over multiple physical or virtual machines.

Variety refers to the source of data. Data can be structured, semi-structured, or unstructured. Structured data can be contained in rows and columns and stored in spreadsheets or relational databases. Although most accounting data is structured, it is estimated that less than 20 percent of all data is structured.

Unstructured data cannot be easily contained in rows and columns and is therefore difficult to search and analyze. Photos, video and audio files, and social media content are examples of unstructured data.

Semi-structured data has characteristics of both structured and unstructured data. It may include some defining details but doesn't completely conform to a rigid structure. For example, the words in an email are unstructured data. The email date and the addresses of the sender and the recipient are structured data. Artificial intelligence algorithms are used to process unstructured and semi-structured data in a way that makes the information useable.

Velocity refers to the speed at which the data is being produced. The amount of data is not only growing; it's growing exponentially as more people gain internet access, and more technology is created that connects humans to machines and machines to machines. Collecting and translating data (especially unstructured data) into usable information is complicated by how quickly new data is generated.

Veracity refers to the quality of the data. Data quality can be negatively affected by untrustworthy data sources, inconsistent or missing data, statistical biases, and human error. The veracity of unstructured data is especially difficult to determine. Machine learning, a type of artificial intelligence based on the idea that systems can learn from data and can identify patterns, is often used to assess data quality.

In summary, a set of data would be considered "big data" if:

- The data set is too large to be managed by traditional methods.
- The data set includes a variety of types of data (structured, semi-structured, and unstructured).
- The amount of data in the data set is expanding rapidly.
- The accuracy and reliability of the data may be uncertain.

Types of Data Analytics

eLectures MBC **LO1** Identify and define the four types of data analytics.

Data analytics can be categorized into four main types, ranging in sophistication from relatively straightforward to very complex. The first category is **descriptive analytics**, which describes what has happened over a given period of time. Simple examples include determining sales trends over a period of time and the relative effectiveness of various social media promotions based on click-through rates. Microsoft Excel and other spreadsheet programs include built-in functions that greatly simplify performing descriptive analytics.

Diagnostic analytics focuses more on why something occurred. This data analytics technique is used to monitor changes in data and often includes a certain amount of hypothesizing: Did the marketing campaign lead to the increase in sales? Did changing the beverage items affect food choices? Did the opening of competing restaurants negatively impact sales growth? Diagnostic analytics is useful because past performance is often a reliable predictor of future outcomes and can greatly aid in planning and forecasting.

Whereas descriptive and diagnostic analytics use data to try to understand what happened and why, **predictive analytics** uses data to try to determine what *will* happen. The movie ***Moneyball*** made the general manager of the **Oakland Athletics**, Billy Beane, famous for using predictive analytics to make personnel decisions in professional baseball. In his evaluation of baseball players, Beane used data to predict player performance so he could assemble the team with the greatest likelihood of winning the World Series. Banks also use predictive analytics to identify and prevent fraudulent transactions by monitoring customer credit card transactions and red flagging those that deviate from a customer behavior profile that was developed from previous transaction and geographic data.

Prescriptive analytics moves beyond what is going to happen to suggesting a course of action for what *should* happen to optimize outcomes. The forecasts created using predictive analytics can be used to make recommendations for future courses of action. For example, if we own a sports bar and determine there is a high likelihood of our local sports team winning the championship this year, we should expand the bar area and add more big-screen televisions to maximize revenues. **Exhibit B.1** summarizes the four types of data analytics.

Exhibit B.1 ■ The Four Types of Data Analytics

Type of Data Analytics	Purpose	Example
Descriptive	To explain what happened	What were sales by month last year?
Diagnostic	To understand why it happened	Did the new advertising campaign cause sales to increase last quarter?
Predictive	To predict what will happen	Does this credit card charge deviate (amount, location, etc.) from past purchases by this credit card holder?
Prescriptive	To determine what should happen	How many servers should be on the schedule for game nights?

Data Analytics in the Accounting Profession

eLectures MBC **LO2** Describe the use of data analytics within the accounting profession.

Accountants are already preparing descriptive analytic reports regularly. Comparative income statements, sales reports by location, inventory valuation reports, and ratio calculations (average collection periods, days' sales in inventory, etc.) are all examples of descriptive analytics.

Budget variance reports and segment reports by region or product line prepared by accountants can be used for diagnostic analytics. Accountants may also work with sales and production managers to analyze the reasons behind changes in operating results. A distributor might want to know how much of the increase in overall sales last year was caused by the transfer of two of its representatives to other sales regions. A grocery store manager might want to know if the winter storm last month

impacted sales in all or just some of the various departments. A production manager might work with the accounting department to determine any correlation between equipment repair costs and the number of units produced over the last two years.

Data analytics should not be limited to only descriptive and diagnostic analysis. Accountants can provide even more value by employing predictive and prescriptive analytics. Accountants can obtain data from a variety of company sources, including enterprise resource planning systems, customer relationship management systems, and point-of-sale systems, to aid them in obtaining insight into future outcomes and providing guidance for future actions. The area of credit granting provides an example. Predictive analytics can help compute credit scores to predict the likelihood of future payments. As a result, prescriptive analytics can aid in suggesting terms for granting credit. Predictive analytics can also be used to help analyze outstanding accounts receivables and determine estimated credit losses based on how much time has elapsed since the credit sale took place.

Many other opportunities exist for accountants to utilize data analytics. Tax accountants can apply data analysis to unique tax issues to suggest optimal tax strategies. Accountants serving as investment advisors can use big data to find patterns in consumer behavior that others can use to build analytic models for identifying investment opportunities.

Perhaps no area of accounting can benefit more from an understanding of data analytics than auditing. Auditors employ data analytics to shift from the sample-based audit model to one based on continuous modeling of much larger data sets. This allows auditors to identify the riskiest areas of an audit by focusing on outliers and exceptions.

The major accounting firms have fully embraced the power of data analytics. **PricewaterhouseCoopers** (PWC), **Deloitte**, **Ernst & Young** (EY), and **KPMG** all devote significant staffing resources to provide data analytics services to their clients. These firms claim they can help their clients optimize their data assets to aid in faster and better decisions. For example, PWC provides a flowchart starting with the building of a data foundation and applies advanced analytics to improving business performance, ultimately leading to opportunities for innovation.

While computers and software are instrumental in the entire process, the human element is the most critical factor in the success of any data analytics program. One commonality among surveys of top company managers is the value placed on data analytics for the company's future. Another commonality is the need for professionals trained in data analytics to help the company attain its goals.

Data Analytics in Accounting

Benford's Law provides an example of how data analytics has been used to uncover fraud in a national call center. Forensic accountants utilized their knowledge of Benford's Law to form evidence of a problem by observing patterns in the data. According to Benford's Law, in any list of financial transactions, the number one should occur as the first digit 30.1 percent of the time, with each successive number occurring as the first digit in lesser percentages, with the number nine occurring less than 5 percent of the time. Forensic accountants examined issued refunds and noticed an excessively high occurrence of the number four. The forensic accountants learned that the company had a policy that required supervisor approval of refunds that exceeded $50. The accountants were able to identify a small group of operators who had been issuing fraudulent refunds to family, friends, and themselves. These fraudulent $40 refunds totaled several hundred thousand dollars.

In order to be useful, data needs to be analyzed. Technology has provided the analyst with powerful tools that allow big data to provide insights that would not have been possible in the past. Still, the most important tool in the analytics toolkit comes from the analyst. Without critical thinking and good judgement, the value would remain locked within the data.

The Analytics Mindset[1]

LO3 Describe the analytics mindset.

The analytics mindset consists of a four-step process of (1) asking the right questions; (2) extracting, transforming, and loading the necessary data; (3) applying appropriate data analytics techniques; and (4) interpreting and presenting the results. **Exhibit B.2** summarizes the steps and requirements of an analytics mindset.

[1] The analytics mindset discussed here is an approach developed by the Ernst & Young Foundation.

Exhibit B.2 ■ Steps of an Analytics Mindset

Steps in the Analytics Mindset	Requirements
Ask the right questions	Understand the objectives of the end user Understand the underlying business processes
Extract, transform, and load the data	Know what to ask for Manage the data security Transform the data into the required format Cleanse the data for completeness and accuracy
Apply the appropriate analytics techniques	Determine if the need is for a confirmatory or an exploratory approach
Interpret and present the results	Use appropriate critical judgement regarding what you see Visually display the results in a format that is easy to understand without unnecessary clutter

Note that while technology is imbedded in this process, the process still begins and ends with the human element of asking the right questions and interpreting the results. Nothing is more critical than the first step of knowing what to ask. The right questions guide the process to find the right data to analyze and interpret.

Asking the right questions requires a few prerequisites. First, you need to know the audience that the analysis is for and what their objectives are. Next, you need to understand the context underlying the problem. For example, to analyze a marketing question you should understand the industry characteristics and the consumer demographics. Without this knowledge you may not select the correct indicators to analyze.

Along with knowing the right questions to ask, an analytical mindset requires you to form an idea of what to expect from the data. For example, when analyzing inventory salability, you would expect to see certain associated movements in sales and receivables.

After your questions are formed, you need to determine the data needed to aid in finding answers to those questions. This requires a knowledge of the data characteristics of the four V's previously mentioned. With this knowledge you can begin the data extraction process. Here you will need to know what data to ask for, how to manage data security, and what form the data will take.

Once you have the data, you will need to transform it into a format suitable for analysis. This is often referred to as data cleaning. Data is rarely found in the form of a nicely organized Excel spreadsheet. Rather, the data will often need to be converted into a proper format and tested for completeness and accuracy. Further, unnecessary data should be removed from the data set.

The data should then be loaded into the proper analysis tool, such as **Tableau** or Microsoft's **Power BI**. Once loaded, the data should again be cleansed to be sure it is ready for analysis in the chosen software.

It is necessary to determine the appropriate technique to analyze the data within the analysis tool. There are a multitude of ways that the data can be analyzed. Possible choices include computing ratios between associated measures, identifying trends among various measures, creating comparisons between dates, and sorting measures. The proper technique to use will be guided by the questions being asked.

In your interpretation of the data, you should ask yourself what do you see and is this what you expected? In other words, do these results make sense or did the results create new questions that require further analysis?

Eventually, the results must be packaged into a presentation that can be shared with the intended audience. Software such as Tableau, Power BI, or Excel can greatly enhance these presentations through their ability to create **visualizations** and **dashboards**. These visualizations can take many forms, from simple tables to bar or pie charts, to more sophisticated scatter plots, map charts, heat maps and more. Dashboards are created by combining multiple visualizations. Interactive dashboards allow users to filter out or drill down on content included in the charts and tables, on demand.

Data Analytic Tools

Technologies used by organizations to analyze data and communicate information to users are known as Business Intelligence (BI) tools. Data warehousing (data storage), data mining (extracting usable insights from data), and reporting and querying software are all BI tools.

Excel and Tableau are two popular BI tools that you will be using in the exercises and problems at the end of this Appendix.

Although Excel and Tableau can be used in similar ways, there are some important differences. Excel is a software application that is used for creating, organizing, and analyzing data. Tableau is a data visualization tool. Although calculations can be performed in Tableau, those calculations are made to create new fields for use in visualizations, not as support for accounting transactions. For example, Excel might be used to calculate sales commission amounts, which are then inputted into the accounting system. Tableau would not be used for that purpose.

Users in both Excel[2] and Tableau can

- Connect with different data sources
- Create visualizations and dashboards
- Work with big data sets

Tableau has much stronger interactivity tools and a more comprehensive selection of chart options. Excel generally has more flexibility and more extensive analytics tools.[3]

Accessing Excel and Tableau

Excel, if not available to you through your school, can be accessed for free by creating a Microsoft account at https://office.live.com/start/Excel.aspx. A free version of Tableau (Tableau Public) is available to you at https://public.tableau.com/en-us/s/. Tableau Public has most of the functions of Tableau Desktop (the full version). However, you can't save your workbooks locally if you're using Tableau Public. Instead, all workbooks are saved online and are accessible to any Tableau user unless you elect to hide your visualizations. Hiding visualizations is done in Settings once you've registered for Tableau on the Tableau website. Walk-through videos are available for every exercise and problem at cambridgepub.com. Tableau tutorial videos are available at https://www.tableau.com/learn/training/.

Data Visualization

eLectures MBC **LO4** Describe data visualization best practices.

As noted above, the final step in the analytics mindset is to present your results. This is often done in the form of a visualization. While it is possible to present results as a bunch of tables full of numbers, visualizations with imagery are often a far better means to convey the raw numbers. Visualizations can be thought of as a blending of the art of design with the science of data.

There is an unlimited number of ways that data can be presented; however, certain best practices exist that can serve as a guide when building a visualization. For example, the exact same data on GDP levels are shown in the three charts in **Exhibit B.3**, but each displays the data differently. The table presents the raw data; however, the reader cannot easily rank the different economies. The two bar charts both show the same data, however the one all in blue makes it far easier to compare economies by showing the data in sorted order. Also, note that adding multiple colors to the other bar chart does nothing to aid the reader, rather it just adds confusion.

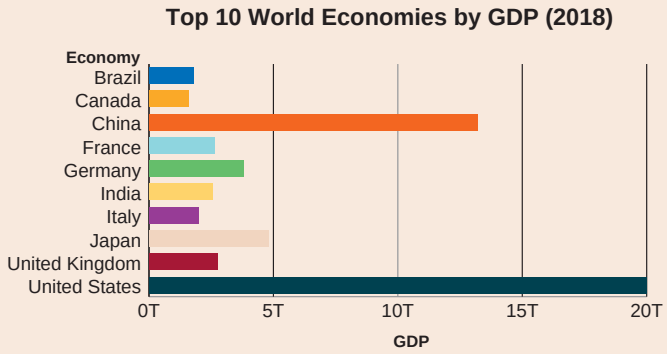
Visualizations can be divided into two primary categories, exploratory and explanatory. **Exploratory visualizations** are meant to allow the reader to explore the data presented in order to do additional analysis. Exploratory visualizations would normally include interactive tools like filters that allow the user to change the level of data displayed. This can be useful when the problem is not clearly defined, and the reader wishes to gain a further understanding of the data.

---

[2] Full functionality in Excel is only available if you have Excel 2010 or newer and you are running a 64-bit version of Windows. To determine the version of Windows on your computer, go to Settings>System>About. The version will be listed in the Device specifications section.

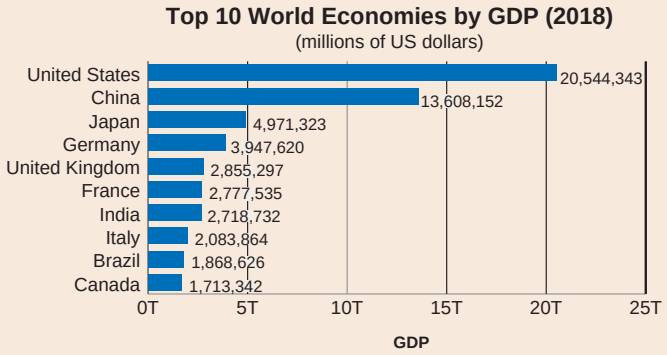
[3] Pan and Blankley, Excel vs. Tableau: See your data differently, *Journal of Accountancy*, February 29, 2020.

Exhibit B.3 ■ Different Displays of the Same Data



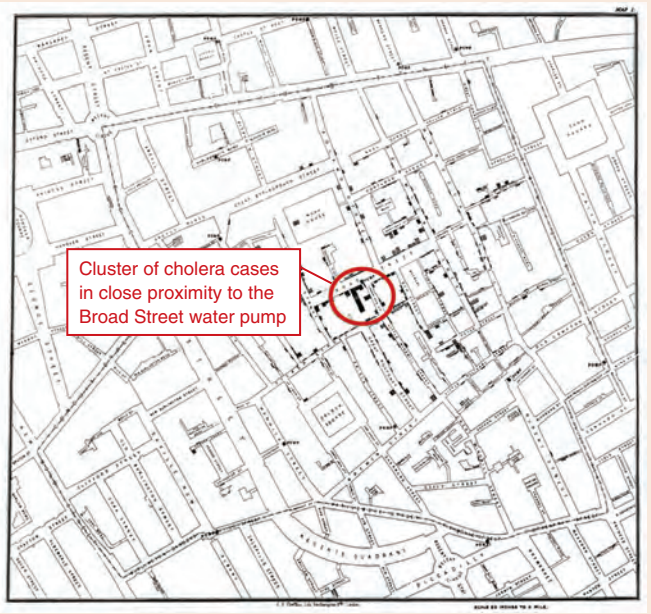
Top 10 World Economies by GDP (2018)
(millions of US dollars)

Economy	
Brazil	1,868,626
Canada	1,713,342
China	13,608,152
France	2,777,535
Germany	3,947,620
India	2,718,732
Italy	2,083,864
Japan	4,971,323
United Kingdom	2,855,297
United States	20,544,343



In contrast to exploratory visualizations, **explanatory visualizations** are used to convey information to the audience. A classic example of such a visualization was prepared in 1854 by the British physician Dr. John Snow. Dr. Snow plotted cholera deaths in central London on a map that also showed the location of water pumps. The visualization identified the relationship between these deaths and the Broad Street water pump and lead to a change in the water and waste systems. Dr. Snow's visualization is shown in **Exhibit B.4**.

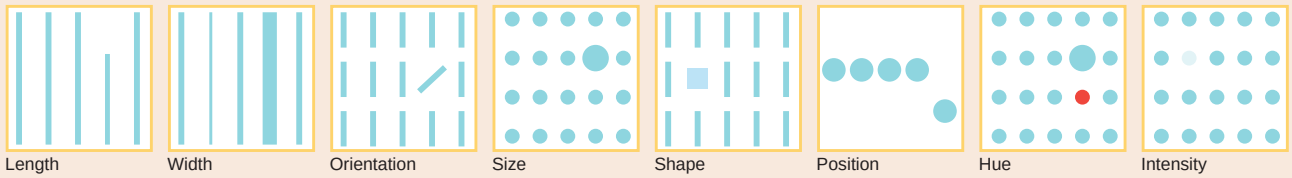
Exhibit B.4 ■ Cholera Deaths in London in 1854



Good visualization design can be enhanced by considering how our brains process visual details such as form, position, and color.

For example, items that are different from the rest become the focus of attention as shown in **Exhibit B.5**. An item that is longer, wider, or in a different orientation will stand out, as will an item that is of a different size, shape, in a different position, or has a different hue or intensity of color.

Exhibit B.5 ■ Displays that Emphasize How Differences Focus Our Attention

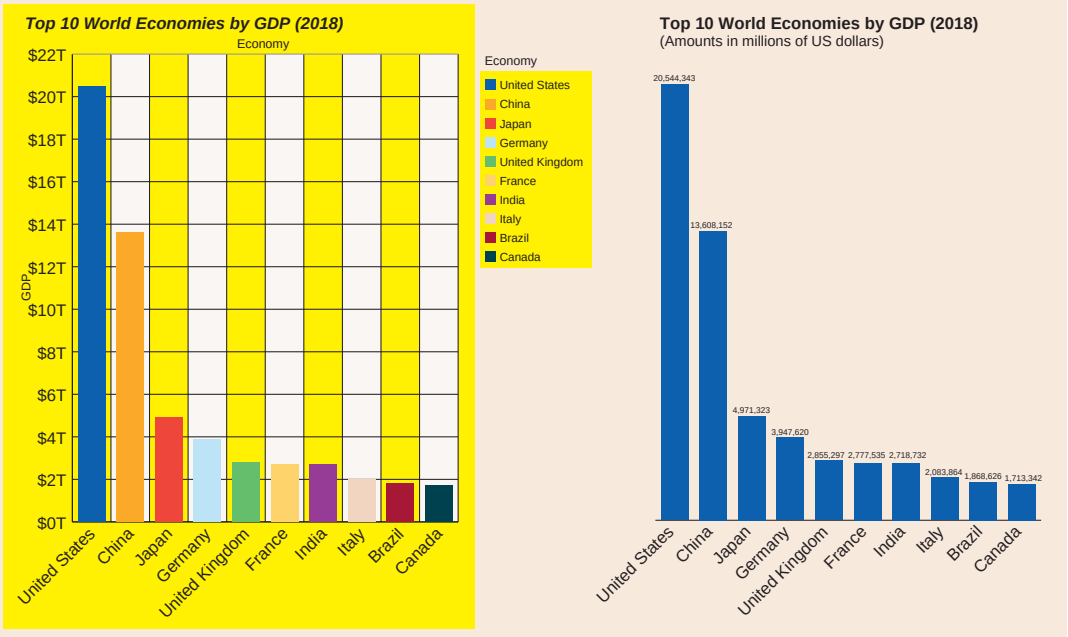


While the use of color can help an item to stand out, it is important to use color correctly. The use of too much color can add to visual clutter. And it's important that color is used consistently, such as always representing a certain year or category. The choice of color is also important since color can convey meanings that differ from one culture to another. For example, red may mean good luck and green may mean jealousy.

Good visualization design requires the removal of items that detract from the message that we are trying to communicate. **Visual clutter** confuses the audience and lessens the chance that they will be able to easily understand the information that is being conveyed. The concept that less is more is the essence of the visualization design principles developed by Edward Tufte, a statistician and professor emeritus at Yale University. Tufte uses the term chart-junk to refer to any unnecessary or confusing elements included in information displays. His principles show that "excellence in statistical graphics consists of ideas communicated with clarity, precision and efficiency."[4]

Exhibit B.6 illustrates **Tufte's principles**. Note in the first visualization all of the visual clutter only serves to distract the audience from seeing the main point that the U.S. is the largest economy based on its GDP. Now notice how much cleaner the second visualization is after removing the distracting yellow background, the color coding of each economy, the redundant labeling, and the unnecessary grid lines.

Exhibit B.6 ■ Illustration of Tufte's Principles



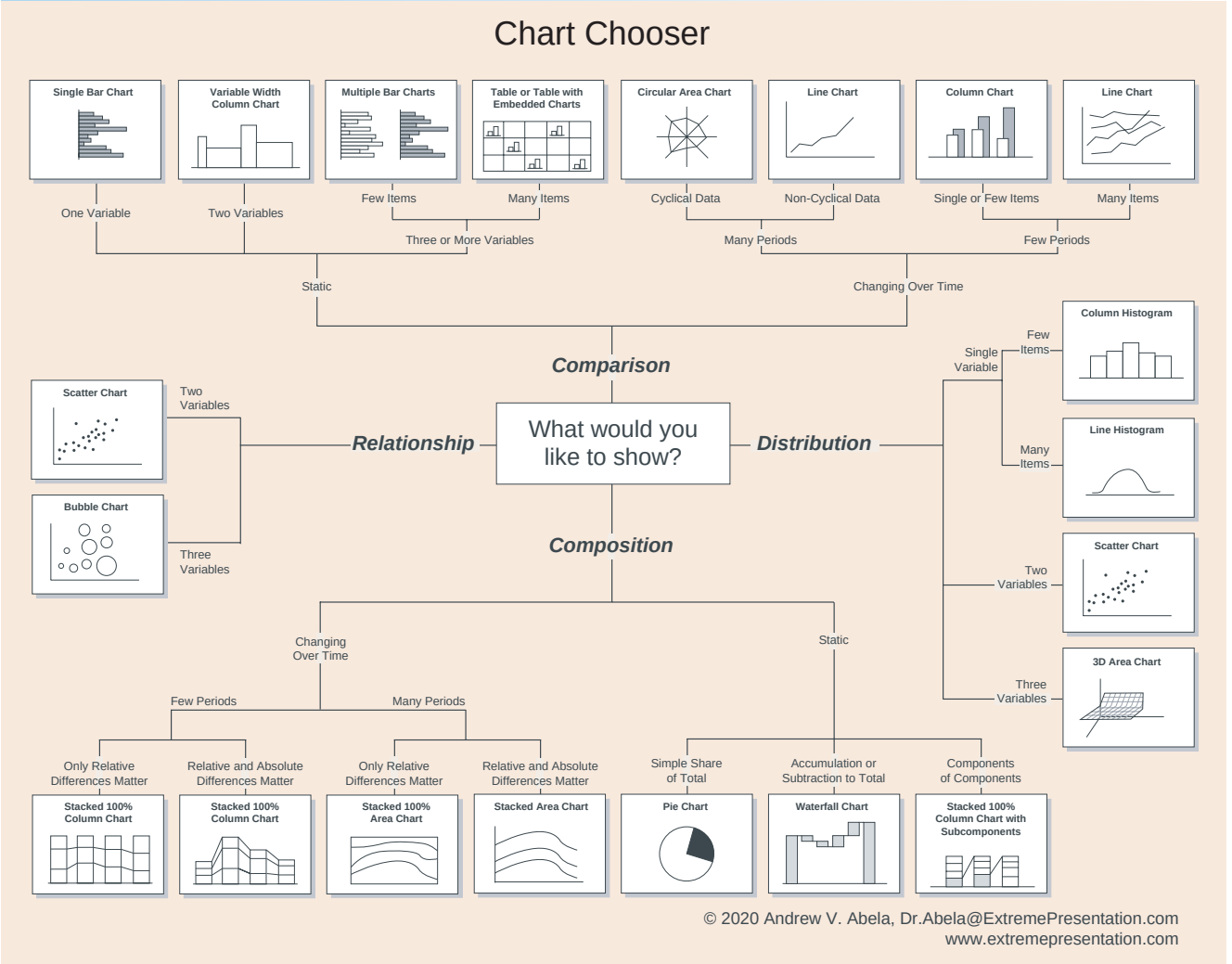
[4] E.R. Tufte, *The Visual Display of Quantitative Information* (Graphics Press, Cheshire, CT 2001).

Good visualization construction also involves choosing the most effective chart type depending on what information is being presented.

The starting point for all of the visualizations we will be discussing is a simple table of data. While the table is excellent for looking up values and can precisely communicate numerical values, visualizations in the form of charts provide the audience an easier method to see what the analyst is attempting to convey.

Among the most used chart types, column and bar charts are best for showing comparisons, line charts are useful for showing trends, pie charts are typically used for showing how individual parts make up a whole, and scatter plots are best for showing relationships and distributions. **Exhibit B.7**, reprinted with permission from the author, provides an excellent tool to help in choosing the correct chart type.[5]

Exhibit B.7 ■ Chart Types



Column (vertical) charts and **bar** (horizontal) charts are best used to compare different categories. Adding labels to the bars rather than just having values showing on the axes makes it easier for the audience to determine these values. Finally, avoid using too many colors that just add to visual clutter.

As a general rule, **line charts** are best for illustrating changes over time and work best with continuous data. Best practices include clearly labeling the axes so the audience knows what is being shown, removing excess clutter such as grid lines and redundant labeling, and avoiding comparing more than five to seven lines.

[5] Abela, Andrew V. (2013). *Advanced Presentations by Design: Creating Communication that Drives Action*. John Wiley & Sons.

Pie charts are best used to show parts of a whole. Be sure the parts add up to 100 percent. Pie charts work best when there are just a few categories. If there are many categories of similar size, consider using a bar or column chart instead. Finally, avoid the temptation to get "fancy" with 3-D imagery and tilting the pie chart.

Scatter plots are useful if the goal is to show correlations between two variables. They are also useful for showing data distributions and clustering, which can identify anomalies and outliers. A **bubble chart** can extend the capability of a scatter plot by adding an additional dimension through changing the size of each bubble in the scatter plot. The more data that is included in a scatter plot or bubble chart, the better are the comparisons that can be made. If the elements being graphed are distributed over a very wide range, the horizontal axis can be converted from a linear to a logarithmic scale (where the numbers on the horizontal axis increase by multiples of a number). Bubble charts should use only circles rather than other shapes. Bubble charts should be scaled based on the area of the circle and not the diameter.

A **map chart** is a good choice if the data being conveyed in the visualization includes geographic locations. Map charts are best at showing relative differences in numerical values among geographic locations rather than precise differences since the values are usually portrayed as differences in a color gradient.

There are several general rules to follow regardless of the chart type. The following list was found from a search of best practices for data visualization charts.[6]

- Time axis. When using time in charts, set it on the horizontal axis. Time should run from left to right. Do not skip values (time periods), even if there are no values.
- Proportional values. The numbers in a chart (displayed as bar, area, bubble, or other physically measured element in the chart) should be directly proportional to the numerical quantities presented.
- Visual clutter. Remove any excess information, lines, colors, and text from a chart that do not add value.
- Sorting. For column and bar charts, to enable easier comparison, sort your data in ascending or descending order by the value, not alphabetically. This applies also to pie charts.
- Legend. You don't need a legend if you have only one data category.
- Labels. Use labels directly on the line, column, bar, pie, etc., whenever possible, to avoid indirect look-up.
- Colors. In any chart, don't use more than six colors.
- Colors. For comparing the same value at different time periods, use the same color in a different intensity (from light to dark).
- Colors. For different categories, use different colors. The most widely used colors are black, white, red, green, blue, and yellow.
- Colors. Keep the same color palette or style for all charts in the series and the same axes and labels for similar charts to make your charts consistent and easy to compare.

Summary of Learning Objectives

Identify and define the four types of data analytics. LO1

- Data analytics can broadly be defined as the process of examining sets of data with the goal of discovering useful information from patterns found in the data.
- Data analytics can be categorized into four types: descriptive, diagnostic, predictive, and prescriptive.

Describe the use of data analytics within the accounting profession. LO2

- Many accountants are already performing descriptive and diagnositc data analytics.
- Accountants can add value by performing predictive and prescriptive data analytics.
- The large accounting firms have devoted large resources to data analytics.
- Being well trained in data analytics is important for future accountants.

[6] https://eazybi.com/blog/data_visualization_and_chart_types/

LO3 **Describe the analytics mindset.**

- Big data has the characteristics of volume, variety, velocity, and veracity.
- Analytics is the process of deriving value from the data.
- An analytics mindset requires critical thinking and judgement.
- The four steps of the analytics mindset include (1) asking the right questions; (2) extracting, transforming, and loading the data; (3) applying the proper analytics techniques; and (4) interpreting and presenting the results.

LO4 **Describe data visualization best practices.**

- Form, position, and color can be used to have elements stand out without any conscious effort by the audience.
- Tufte's principles of design emphasize the elimination of visual clutter that serves to distract from the ability of a visualization to convey its message.
- Use of the proper chart type can help the intended audience to visualize comparisons, compositions, distributions, and relationships in the data.

Key Concepts and Terms

Bar (p. 526)
Benford's Law (p. 521)
Big data (p. 519)
Bubble chart (p. 527)
Column chart (p. 526)
Dashboards (p. 522)
Data analytics (p. 519)
Descriptive analytics (p. 520)
Diagnostic analytics (p. 520)
Explanatory visualizations (p. 524)
Exploratory visualizations (p. 523)
Line charts (p. 526)
Map chart (p. 527)
Pie chart (p. 527)
Predictive analytics (p. 520)
Prescriptive analytics (p. 520)
Scatter plot (p. 527)
Tufte's principles (p. 525)
Variety (p. 519)
Velocity (p. 519)
Veracity (p. 519)
Visual clutter (p. 525)
Visualizations (p. 522)
Volume (p. 519)

Questions

LO1, 2 **QB-1.** Which of the following are the four categories of data analytics?

Homework MBC

a. Descriptive, diagnostic, predictive, prescriptive
b. Expressive, diagnostic, predictive, prescriptive
c. Descriptive, analytical, predictive, prescriptive
d. Descriptive, diagnostic, prognostic, prescriptive

LO3 **QB-2.** Which of the following are four characteristics of big data?

Homework MBC

a. Volume, variety, vagueness, veracity
b. Volume, variety, velocity, veracity
c. Volume, validate, velocity, veracity
d. Volume, variety, velocity, vulnerability

LO3 **QB-3.** What is the correct order of the steps in the analytics mindset?

Homework MBC

a. Extract, transform, and load the data; ask the right questions; apply the proper analytics techniques; interpret and present the results.
b. Ask the right questions; extract, transform, and load the data; apply the proper analytics techniques; interpret and present the results.
c. Ask the right questions; extract, transform, and load the data; interpret and present the results; apply the proper analytics techniques.
d. Ask the right questions; apply the proper analytics techniques; extract, transform, and load the data; interpret and present the results.

QB-4. Charts are used in visualizations to convey the following primary types of information: LO4

a. comparisons, compositions, distributions, and relationships.

b. comparisons, historical, distributions, and relationships.

c. comparisons, compositions, forecasts, and relationships.

d. geographical, compositions, distributions, and relationships.

QB-5. Which of the following statements is not true regarding the use of color in a chart? LO4

a. Use at most six different colors in a chart.

b. To show changes in an item over time use a color gradient rather than different colors.

c. Always use color in a chart to differentiate items.

d. Use the same color palette in a chart series.

Assignments with the MBC logo in the margin are available in myBusinessCourse. See the Preface of the book for details.

Exercises

EB-6. **Public Accounting Firms and Data Analytics** LO1, 2

Go to PWC.com and select "Services" and then "Data and Analytics." Choose a topic and write about how PWC is using data analytics to help its clients.

EB-7. **Public Accounting Firms and Data Analytics** LO1, 2

Go to KPMG.com and select "Insights." Under "Areas of interest," select "Special Attention" and then "Data and Analytics." Choose a topic and write about how KPMG is using data analytics to help its clients.

EB-8. **Public Accounting Firms and Data Analytics** LO1, 2

Go to Deloitte.com and select "Services" and then "Analytics." Choose a topic and write about how Deloitte is using data analytics to help its clients.

EB-9. **Public Accounting Firms and Data Analytics** LO1, 2

Go to EY.com and enter Big data and analytics in the search bar. Choose a topic and write about how Ernst & Young is using data analytics to help its clients.

Problems

Problems PB-10 through PB-12 use financial statement data for S&P 500 companies for the years 2015 through 2019. The Excel file Compustat_SP500_2015_2019.xlsx is accessible on the textbook's website. A video demonstrating Tableau tools used to answer the questions in the next two problems is also available on the website.

PB-10. **Building Basic Tableau Visualizations** LO3, 4

a. Connect the Tableau software to the Excel file Compustat_SP500_2015_2019.xlsx. This file consists of four workbooks. First bring in the Balance sheet workbook and then join both the cash flow statement and the income statement workbook to the balance sheet workbook using both of the fields company name and year.

b. What is the sum of net income for all firms in the database for all years combined? One way to determine this is to drag the measure Net income to the canvas.

c. How many unique companies are included in the database? One way to determine this is to drag the dimension Company name to the rows shelf and then select Measure Count(Distinct) from the pull down menu on the Company name pill.

d. How many distinct firms are there in each segment? One way to determine this is to drag the dimension Segment to the Columns shelf in the visualization created in part *c*. The totals for each segment will appear if the Show marks label is checked in the Label card.

e. What is the sum of total assets for all companies in each segment for the year 2018? One way to determine this is to drag the dimension Segment to the Columns shelf and then drag the Total Assets measure to the Rows shelf. Next drag the Year dimension to the filters shelf, select year as the filter, click next and then check 2018. Totals for total assets can be seen in the tool tip by hovering over any bar or by checking Show marks label in the Label card.

f. What firm had the most sales in 2018? What segment was this firm in? One way to determine this is to drag the dimension Company name to the rows shelf and drag the measure Sales to the columns shelf. Next drag the Year dimension to the filters shelf, select year as the filter, click next and then check 2018. Segments can be highlighted by dragging the dimension Segment over the color card. Finally sort the Company names by Sales by clicking the sort icon in the tool bar.

g. Save the file for future use.

LO3, 4

PB-11. Tableau Visualizations to Analyze Accounting Performance Measures

Homework MBC

You recently joined a firm as a junior financial analyst, and you would like to make a good impression by showing your manager the power of visualizations for analyzing data. In order to get a feel for the Tableau software and the dataset you created of financial statement data for S&P 500 firms, you decided to create a few very basic visualizations. Two widely used ratios to analyze company performance are gross profit percentage and return on sales. You decide to create a visualization that compares these two ratios by segment and further compares segment performance to the median values of these ratios for the entire database of companies.

a. Because of the way cost of goods sold is reported for companies in the real estate segment you decide to exclude this segment from the visualization. After excluding real estate, for the year 2017, which segment reported the highest median value for gross profit percentage and for return on sales?

b. Did any segment report a higher median return on sales than the upper band of the 95 percent confidence interval of overall median return on sales in 2015?

c. What company had the highest gross profit percentage in 2018 for the segment with the highest median gross profit percentage?

LO3, 4

PB-12. Using Tableau to Analyze Inventory

Homework MBC

You have learned of the importance of a company being able to sell its products in a timely fashion, and that the ratio of days sales in inventory provides this useful information. You decide a dashboard would be helpful in seeing if this ratio is improving or declining in the consumer discretionary and the consumer staples segments between 2017 and 2018. You build two sheets that are included in the dashboard. The first sheet shows the level of the ratio for each segment for the two years in question. The second sheet shows the change in the ratio between the two years.

Has the ratio days sales in inventory improved or declined in the consumer discretionary and the consumer staples segments between 2017 and 2018. By how much?

LO2, 3

PB-13. Using Tableau for Fraud Detection

Homework MBC

Benford's Law represents a powerful tool in the forensic accountant's toolkit to aid in the detection of fraud. Benford's Law is a mathematical law that recognizes the leading (first) digit in many real-life number sets is distributed in a certain manner, and often not in the manner that a fraudster would expect. Specifically the number 1 occurs as the first digit approximately 30 percent of the time, with each succeeding digit appearing less often as follows: 1–30%, 2–18%, 3–12%, 4–10%, 5–8%, 6–7%, 7–6%, 8–5%, and 9–5%. Fraudsters who are unaware of this natural ordering will often arrange digits in a random order that deviates from Benford's Law.

In Part A of this problem you will use Tableau to show how a natural data set of GDP by country conforms to Benford's Law and how a random set of numbers does not. In Part B you will use the same data used in an actual court case to convict a fraudster of embezzlement. Finally, in Part C you will use Benford's Law to test a new reimbursement procedure for possible fraud. A video demonstrating the Tableau tools used in this problem is available on this textbook's website.

Part A Use Tableau to show how a natural data set of GDP by country conforms to Benford's Law and how a random set of numbers does not.

- Download the file **GDP Tableau.xlsx** from the textbook website. The file contains World Bank GDP data by country for 2018, along with a separate column of random numbers that was generated in Excel using the command `=RAND()*1000`.

- After you have uploaded the workbook to Tableau, create two calculated fields.
- The first calculated field will pull the first digit from each country's GDP amount. Choose Analysis > Create Calculated Field and name the calculation First Integer. Then either type or paste the following formula in the formula area: `LEFT(STR([GDP]),1)`
- Next create a second calculated field named Benfords Law by typing or pasting the following in the formula area: `LOG(INT([First Integer])+1)- LOG(INT([First Integer]))`
- To create the visualization, drag First Integer from the Dimensions area to Columns and drag Number of Records from the Measures area to Rows. Click Sum(Number of Records) on Rows to show the pull-down menu and choose Quick Table Calculation > Percent of Total. The visualization should now show a bar chart with the bars conforming to Benford's Law.
- While it is relatively easy to see that the data conforms to Benford's Law, with a little more work the visualization can be significantly enhanced. To do this, drag Benfords Law from the Measures area of the Data pane to Detail on the Marks card, and then click Benfords Law on the Marks card and choose Measure > Minimum.
- Next switch from the current Data pane to the Analytics pane and then drag Distribution Band over the chart and drop it on the cell icon in the pop-up. A dialog box will appear. Under computation change the value to percentages of 90,100,110 and select Percent of to be Min(Benfords Law). Choose a fill line as the thick black line and then click OK.
- Finally click on the Label icon in the Marks section and select the Show marks labels box.

a. Does the GDP data appear to conform to Benford's Law?

Now return to the Data pane and create a new calculated field for the random numbers by naming the calculation Random Values and typing or pasting the following formula in the formula area: `LEFT(STR([Random]),1)`

- Drag the Min(Benfords Law) pill out of the Marks area to remove the bands and drag Random Values from the Dimensions area on top of First Integer to replace it in the visualization. If both pills remain in the columns section, simply drag First Integer away.

b. Do the random values appear to conform with Benford's Law?

Part B Use the same data used in an actual court case to convict a fraudster of embezzlement.

In the 1993 court case *State of Arizona v. Wayne James Nelson* Benford's Law was used to convict the defendant of defrauding the state of nearly $2 million by diverting money to a nonexistent vendor. Nelson tried to make the checks appear random; however, he was unaware that these check amounts should actually follow Benford's Law much closer than the random distribution he created. Download the file **Arizona fraud.xlsx** from the textbook website and follow the same procedure as you did in Part A above.

a. From a casual observation of the checks, can you detect anything suspect?
b. After using Benford's Law, does the list of checks appear suspect?

Part C Use Benford's Law to test a new reimbursement procedure for possible fraud.

Wally's Enterprises has been reimbursing its employees for business expenses after the employee submits detailed evidence of the expense, such as paid receipts. Management has recently changed the reimbursement policy because of the time spent checking all the submitted evidence, with an especially high volume of smaller reimbursement requests. The new policy only requires evidence be submitted if the reimbursement request exceeds $50. As the company's internal auditor, you are concerned that this policy change may result in fraudulent reimbursement requests. In order to test the new policy, you have gathered a random sample of 100 reimbursement requests from both before and after the policy change. This data is located in the file **Expense Reimbursement.xlsx** on the textbook's web page. Download this file and use Benford's Law to test whether the new policy appears to have resulted in any fraud.

a. Do the reimbursement requests prior to the policy change appear to follow Benford's Law?
b. Do the reimbursement requests occurring after the policy change appear to follow Benford's Law?
c. What, if anything, leads you to believe that fraud may be occurring?

LO3, 4

PB-14. Segment Reports Using Tableau (Descriptive and Diagnostic Analytics)

Southern Comforts, Inc. is a department store chain with stores in North Carolina, Tennessee, Kentucky, and West Virginia. Its corporate headquarters are located in Charlotte, North Carolina.

In the past, the store owners only received financial reports for the company operations overall. They have recently asked for reports of costs and profitability by segment (Location and department). Southern Comforts' locations include the four stores (Charlotte, Nashville, Virginia Beach, and Louisville) and the corporate office (Charlotte HQ). Departments include the product lines (Mens, Womens, Kids, Shoes, and Home) and the overhead expense types (Facilities, Labor, and Other).

They have provided you with an Excel workbook that includes Southern Comfort transactions for 2020. (The workbook, **Segment Report Data Set Tableau.xlsx**, is accessible on the textbook's website. A video demonstrating Tableau tools used to answer the questions in this problem is also available on the website.)

The first step is to make sure the data is in the form needed.

- Convert the data to a table.
- Check each column to make sure there is no missing or inconsistent data. Make any corrections necessary. *HINT:* There are two errors. One error is in the Month column. The other is in the State column.
- All transactions are included in the Transactions column. You will need to separate revenue transactions from expense transactions. Use the IF function to create a Revenues column and an Expenses column. (All positive numbers in the Transactions columns are Revenues; all negative numbers are expenses. You may want to change the sign of the amounts in the Expense column to positives.)
- Add a column after Month and call it Month Name. Use the TEXT function to convert the date format to a text format (name of the month).

Save the file with a new name. Open Tableau and import the workbook.

1. Create Sheets in Tableau to answer the following questions:
   a. Which store was the most profitable (in dollars)? What was that store's profit? Which store had the most revenue? How much?
   b. Which month had the highest revenue? How much? What percentage of total sales occurred during that month? (Round % to two decimals.) Which month was the least **profitable** (in dollars)? What was the net profit that month? *HINT:* You will need to include the corporate costs to determine net profit.
   c. What was the total gross margin (in dollars) for 2020? Which store had the highest gross margin? How much? *HINT:* You will need to filter out the overhead expense categories (Facilities, Labor, and Other).
   d. What was the total gross profit ratio for 2020? (Round % to two decimals.) Which store had the lowest gross profit ratio? Which product line (department) had the highest gross profit ratio?
2. The store with the highest sales (b.) and the highest gross margin (d.) was not the most profitable (a.). Why? Look at the revenues, gross profit margins and ratios, and the overhead expenses for both stores.
3. Create charts on your Sheets and use them to create an interactive dashboard. Include filters to allow users to look at selected data.

LO3, 4

PB-15. Job Profitability Using Tableau (Descriptive and Diagnostic Analytics)

Harvard Products is a job shop (a company that manufactures custom products in small batches). Each batch is managed by one of Harvard's four project managers. Manufacturing facilities are located in Illinois, Wisconsin, Michigan, and Indiana.

The President of Harvard Products has asked for information about costs and profits by job, Location, customer, and project manager. A summary of costs (by job) is included in the **Job Order Data Set-Tableau.xlsx** file available on the textbook's website. A video demonstrating Tableau tools used to answer the questions in this problem is also available on the website.

Before uploading the workbook to Tableau, add columns to the Data sheet to separate out account amounts (revenue, direct material, etc.). The IF function is useful here.

Once you've uploaded the workbook to Tableau, change the data type for Job # to text. (It should be a Dimension, not a Measure.) Give Location a geographic role.

Create sheets in Tableau to answer the following questions:

1. Which customer was the most profitable for Harvard? What was the job number for the most profitable job for that customer? What was the average revenue on jobs for that customer ? What was the average revenue for all jobs?
2. Which Location was the least profitable (in dollars) for Harvard? What appears to have contributed more to the lower profits at that Location—size of jobs (average revenue) or profit margin ratios? Which Location had the highest average revenue per job?

HINT: To answer the next three questions, you will need to create four new fields. One for the gross profit ratio and one for the three expense ratios.

3. Which Location had the highest gross profit ratio? Which cost element had the highest expense ratio in that Location? What Location had the lowest gross profit ratio? Which cost element had the highest expense ratio in that Location? Which cost element had the largest expense ratio difference between Locations? What factors might be causing those differences?
4. The President of Harvard Products has decided to give performance bonuses to project managers. Determine which project manager will receive the highest bonus if performance is based on:
   a. Average revenue per job?
   b. Number of jobs? *HINT:* You'll need to duplicate Job # to get a count. (Counts would be Measures.)
   c. Total profit?
   d. Profit margin ratios?
5. Create a Dashboard the president could use when evaluating different bonus calculation performance measures. Use a map for one of the sheets included in the dashboard.
6. Based solely on the information you have available, would you encourage management to close the facility with the lowest profit margin? Why or why not? Include in your answer the information you might need to have before making a final decision.

PB-16. Determining Fixed vs. Variable Cost Components Using Tableau (Diagnostic Analytics) LO3, 4

Genessee Industries introduced a new product last year (6582-D). Although it was very popular, it wasn't very profitable. Management has asked you to provide them with information to help them set a new sales price.

You know that the direct costs per unit are $25 for direct materials and $5 for direct labor. You are given information about last year's monthly production levels and manufacturing overhead costs (indirect materials, indirect labor, and other). (That information is included in the **Fixed and Variable Data Set Tableau.xlsx** file available on the textbook's website. A video demonstrating Tableau tools used to answer the questions in this problem is also available on the website.)

1. Use the Trendline tool in Tableau to determine the fixed and variable components for each of the three manufacturing overhead components. Use the default model type (linear). (Round all amounts to two decimals.)
   a. What is the cost formula for indirect materials?
   b. What is the cost formula for indirect labor?
   c. What is the cost formula for other manufacturing overhead?
2. What is the cost formula for 6582-D (per month)?
3. Determine the minimum sales price Genessee could charge to achieve a monthly gross profit of $7,500 next year. Management expects unit sales will average 1,500 per month. *HINT:* You won't be able to use Tableau for this. Start by determining the CVP formula.

PB-17. Forecasting Using Tableau (Predictive Analytics) LO3, 4

Melton Manufacturing opened in January 2019. Sales have increased significantly in the first two years of operations, and management is now looking to expand production capacity. To finance the purchase of a new factory, they would need to either raise capital or borrow funds.

They have asked you to make some projections for the next year of operations. They intend to share these with potential investors and lenders.

Information about unit sales, sales revenues, and net profits for the past two years is included in the **Forecasting Data Set Tableau.xlsx** file available on the textbook's website. (A video demonstrating Tableau tools used to answer the questions in this problem is also available on the website.)

1. Create a worksheet that shows revenues, net operating income, and units sold by month. Add a Forecast. *HINT:* There will be three graphs
   a. Extend the trendline out for one full year. Since you have complete data for each month, do not ignore the last month of 2020.
   b. Use a prediction interval of 95%

2. Use the forecasts to determine:
   a. Expected unit sales in October 2021
   b. Expected sales revenue in June 2021
   c. Expected net operating income in December 2021
3. Open the Describe Forecast screen. Use the Summary tab to determine:
   a. What percent of the forecast was attributed to seasonality for net operating income?
   b. What was the range given for number of units sold for January 2021?
5. Duplicate the worksheet as a cross-tab and swap the axes. Filter out details for net operating income and units sold. For December 2021, what is Tableau's:
   a. estimate?
   b. lower prediction interval?
   c. upper prediction interval?
6. How could Melton use forecasts to manage or expand the business?

A video demonstrating the use of Microsoft Excel to use Benford's Law in the detection of fraud is available on the textbook's website to assist in solving problem PB-18.

LO2, 3

PB-18. Using Excel for Fraud Detection

Benford's Law represents a powerful tool in the forensic accountant's toolkit to aid in the detection of fraud. Benford's Law is a mathematical law that recognizes the leading (first) digit in many real-life number sets is distributed in a certain manner, and often not in the manner that a fraudster would expect. Specifically the number 1 occurs as the first digit approximately 30 percent of the time, with each succeeding digit appearing less often as follows: 1–30%, 2–18%, 3–12%, 4–10%, 5–8%, 6–7%, 7–6%, 8–5%, and 9–5%. Fraudsters who are unaware of this natural ordering will often arrange digits in a random order that deviates from Benford's Law.

In Part A of this problem you will use Microsoft Excel to show how a natural data set of GDP by country conforms to Benford's Law and how a random set of numbers does not. In Part B you will use the same data used in an actual court case to convict a fraudster of embezzlement. Finally, in Part C you will use Benford's Law to test a new reimbursement procedure for possible fraud. A video demonstrating the Excel tools used in this problem is available on the textbook's website.

Part A Use Microsoft Excel to show how a natural data set of GDP by country conforms to Benford's Law and how a random set of numbers does not.

- Download the file **GDP.xlsx** from the textbook website. The file contains World Bank GDP data by country for 2018.
- In order to use Benford's Law you need to first extract the leading digit from each country's GDP amount. To do this, place the cursor in cell C2 and input the formula `=Left(B2,1)`. Copy this formula down column C for each country.
- Next in cells F2 through F10 input the numbers 1 through 9. In cell G2 input the formula `=COUNTIF($c$2:$C$205,F2)` and copy the formula down for each number 1 through 9. This formula goes through the entire range of extracted first digits in column C and records the count of these digits in the cell if it matches the number in column F.
- Sum the column total in cell G11.
- Next determine the percentage that each leading digit appears by dividing the amount in column G by the total of these amounts in cell G11 and place this figure in column H.
- In column I, compute the predicted occurrences of each digit (given above) by placing the formula `=Log10(1/F2+1)` in cell I2 and copying the formula down the column.
- Finally create a Combo chart to visualize these results by highlighting cells H1:I10 and selecting Combo chart.

a. Do the naturally occurring GDP amounts appear to follow Benford's Law?

- Next replace the GDP amounts with random numbers to see if random numbers also obey Benford's Law.
- Input the formula `=Rand()*1000` in cell B2 and copy this formula down the column.
- Observe the results in the table and the chart. Try to recalculate the spreadsheet several times to obtain different sets of random numbers.

b. Do random numbers appear to follow Benford's Law?

Part B Use the same data used in an actual court case to convict a fraudster of embezzlement.

In the 1993 court case *State of Arizona v. Wayne James Nelson* Benford's Law was used to convict the defendant of defrauding the state of nearly $2 million by diverting money to a nonexistent vendor. Nelson tried to make the checks appear random; however, he was unaware that these check amounts should actually follow Benford's Law much closer than the distribution he created. Download the file **Arizona fraud.xlsx** from the textbook website and follow the same procedure as you did in Part A above.

a. From a casual observation of the checks, can you detect anything suspect?
b. After using Benford's Law, does the list of checks appear suspect?

Part C Use Benford's Law to test a new reimbursement procedure for possible fraud.

Jimmy's Enterprises has been reimbursing its employees for business expenses after the employee submits detailed evidence of the expense, such as paid receipts. Management has recently changed the reimbursement policy because of the time spent checking all the submitted evidence, with an especially high volume of smaller reimbursement requests. The new policy requires evidence be submitted only if the reimbursement request exceeds $50. As the company's internal auditor, you are concerned that this policy change may result in fraudulent reimbursement requests. In order to test the new policy, you have gathered a random sample of 100 reimbursement requests from both before and after the policy change. This data is located in the file **Expense Reimbursement Excel.xlsx** on the textbook's web page. Download this file and use Benford's Law to test whether the new policy appears to have resulted in any fraud.

a. Do the reimbursement requests prior to the policy change appear to follow Benford's Law?
b. Do the reimbursement requests occurring after the policy change appear to follow Benford's Law?
c. What, if anything, leads you to believe that fraud may be occurring?

PB-19. Segment Reports Using Excel (Descriptive and Diagnostic Analytics) LO3, 4

Southern Comforts, Inc. is a department store chain with stores in North Carolina, Tennessee, Kentucky, and West Virginia. Its corporate headquarters are located in Charlotte, North Carolina.

In the past, the store owners only received financial reports for the company operations overall. They have recently asked for reports of costs and profitability by segment (location and department). Southern Comforts' locations include the four stores (Charlotte, Nashville, Virginia Beach, and Louisville) and the corporate office (Charlotte HQ). Departments include the product lines (Mens, Womens, Kids, Shoes, and Home) and the overhead expense types (Facilities, Labor, and Other).

They have provided you with an Excel workbook that includes Southern Comfort transactions for 2020. (The workbook, **Segment Report Data Set.xlsx**, is accessible on the textbook's website. A video demonstrating Excel tools used to answer the questions in this problem is also available on the website.).

The first step is to make sure the data is in the form needed.

- All transactions are included in the Transactions column. You will need to separate revenue transactions from expense transactions. Add two columns to the table. Use the IF function to create a Revenues column and an Expenses column. (All positive numbers in the Transactions columns are Revenues; all negative numbers are expenses.) *HINT:* To save some time, convert the data to a Table.
- Add a column after Month and call it Month Name. Use the TEXT function to convert the date format to a text format.

1. Use pivot tables to answer the following questions:
   a. Which store was the most profitable (in dollars)? What was the store's profit?
   b. Which store had the most revenue? How much? Which month had the highest revenue? How much? What percentage of total sales occurred during that month? (Round % to two decimals.)
   c. Which month was the least **profitable** (in dollars)? What was the net profit that month?
   d. What was the total gross margin (in dollars) for 2020? Which store had the highest gross margin? How much? *HINT:* Consider creating a calculated field using the Revenues and Expenses columns in the data sheet. Slicers can be used to filter out the overhead expense categories (Facilities, Labor, and Other).
   e. What was the total gross profit ratio for 2020? (Round % to two decimals.) *HINT:* You can create another calculated field using the gross margin field from d. Which store had the lowest gross profit ratio? What was it? Which product line (department) had the highest gross profit ratio? What was it?

2. The store with the highest sales (b.) and the highest gross margin (d.) was not the most profitable (a.). Why? Look at the revenues, gross profit margins and ratios, and the overhead expenses for both stores.
3. Create pivot charts from some of your pivot tables and use them to create an interactive dashboard. Include slicers on the dashboard that allow management to filter by location or month.

LO3, 4 **PB-20. Job Profitability Using Excel (Descriptive and Diagnostic Analytics)**

Harvard Products is a job shop (a company that manufactures custom products in small batches). Each batch is managed by one of Harvard's four project managers. Manufacturing facilities are located in Illinois, Wisconsin, Michigan, and Indiana.

The President of Harvard Products has asked for information about costs and profits by job, location, customer, and project manager. A summary of costs (by job) is included in the **Job Order Data Set.xlsx** file available on the textbook's website. A video demonstrating Excel tools used to answer the questions in this problem is also available on the website. *HINT:* Add columns to the Data sheet to separate out account amounts (revenue, direct material, etc.). The IF function is useful here.

Create PivotTables to answer the following questions:

1. Which customer was the most profitable for Harvard? What was the job number for the most profitable job for that customer? What was the average revenue on jobs for that customer? What was the average revenue for all jobs?
2. Which location was the least profitable (in dollars) for Harvard? What appears to have contributed more to the lower profits at that location – size of jobs (average revenue) or profit margin ratios? Which location had the highest average revenue per job?
3. The President of Harvard Products has decided to give performance bonuses to project managers. Determine which project manager will receive the highest bonus if performance is based on:
   a. Average revenue per job?
   b. Number of jobs?
   c. Total profit?
   d. Profit margin ratios?
4. Based solely on the information you have available, would you encourage management to close the facility with the lowest profit margin? Why or why not? Include in your answer the information you might need to have before making a final decision.

LO3, 4 **PB-21. Determining Fixed vs. Variable Cost Components Using Excel (Diagnostic Analytics)**

Genessee Industries introduced a new product last year (6582-D). Although it was very popular, it wasn't very profitable. Management has asked you to provide them with information to help them set a sales price that will provide them with a monthly gross profit of $7,500. (Last year's sales price was $75 per unit.)

You know that the direct costs per unit are $25 for direct materials and $5 for direct labor. You are given information about last year's monthly production levels and manufacturing overhead costs (indirect materials, indirect labor, and other). (That information is included in the **Fixed and Variable Data Set.xlsx** file available on the textbook's website. A video demonstrating Excel tools used to answer the questions in this problem is also available on the website.)

1. Determine the cost formula for 6582-D. Use Excel's regression analysis tool to calculate the fixed and variable manufacturing overhead costs. Round elements to two decimal places.
   a. Check the 95% confidence level. Consider checking the box to add a line fit plot for each indirect cost element to show the relationship in chart form. (You may need to change the minimum bound on the horizontal axis to 700 to see the line clearly.)
2. Use the prior year data to create a graph of the various overhead costs by month.
   a. Create a combo chart as follows:
      i. The primary vertical axis in dollars, and the secondary vertical axis is units of production.
      ii. The horizontal axis is Months.
      iii. Units produced should be a clustered column type; the overhead cost elements should be line type. (Units produced would be linked to the secondary vertical axis.
3. Use Excel's Goal Seek tool to determine the sales price required to meet the $7,500 gross profit goal. Management believes monthly sales will average 1,500 next year. Assume the company will not maintain any inventory of finished goods.
4. Discuss how Goal Seek (or any other Excel tool) might help management with CVP Analysis.

PB-22. Forecasting Using Excel (Predictive Analytics) LO3, 4

Melton Manufacturing opened in January 2019. Sales have increased significantly in the first two years of operations, and management is now looking to expand production capacity. To finance the purchase of a new factory, they would need to either raise capital or borrow funds.

They have asked you to make some projections for the next year of operations. They intend to share these with potential investors and lenders.

Information about unit sales, sales revenues, and net profits for the past two years is included in the **Forecasting Data Set.xlsx** file available on the textbook's website. A video demonstrating Excel tools used to answer the questions in this problem is also available on the website.

1. Create three line graphs in Excel (one for units sold, one for sales revenue, and one for net operating income). Add trendlines to all graphs.
   a. Extend the trendline out for 12 months.
   b. Use the Polynomial (Order 2) trendline option for all charts
   c. To see how closely the trendline matches the data, check the *Display R-squared value on chart* box. The closer the R-square value is to 1, the better the match.
2. Create the same three graphs using the Forecast Sheet tool (line charts) in Excel.
   a. Set the *Forecast End* to 12/1/2021.
   b. Use an 85% *Confidence Interval.*
   c. Check the *Include forecast statistics* box.
   d. Leave remaining defaults as is.
3. Use the trendline graphs to determine: (*HINT:* To help identify the answers, display gridlines. Consider changing vertical axis bounds.)
   a. Expected unit sales in October 2021
   b. Expected sales revenue in June 2021
   c. Expected net profits in December 2021
4. Use the Forecast sheets to determine:
   a. Range of expected unit sales in October 2021 (Upper to lower Confidence bounds)
   b. Expected sales revenue in June 2021 (Upper to lower Confidence bounds)
   c. Range of expected net profits in December 2021 (Upper to lower Confidence bounds)
5. To evaluate the Forecast sheets, rerun the forecasts. This time change the *Forecast Start* date to 1/1/2020 to see what the model would have predicted for 2020. (Leave all other options the same as B. above.) Were the predictions higher or lower than the actual results? What could have caused the differences?

PB-23. Utilization of Constrained Resources Using Excel (Prescriptive analytics) LO3

Backyard Helpers, Inc. is a small manufacturing company with 18 different gardening tools in its product line. All of the products are fabricated using the same equipment.

Recently, sales demand has increased. Unfortunately, Backyard Helpers cannot produce enough products with existing equipment to meet that demand. Facilities can be expanded, and new equipment purchased, but it will be at least two years before that happens. The production manager needs to make production scheduling decisions now.

The **Constrained Resource Data Set.xlsx** file available on the textbook's website includes information about demand, sales price, cost, and fabrication time on the shared equipment for each of Backyard Helpers' products. A video demonstrating Excel tools used to answer the questions in this problem is also available at cambridgepub.com.

Maximum machine time is 40,500 minutes per month. The demand for all products is spread equally throughout the month. Fixed costs (manufacturing, selling, and administrative) total $755,750 per month. Backyard Helpers maintains no inventory of finished goods. (All units produced are sold during the month.)

1. Ignore machine time limits in answering the following:
   a. Which product has the highest contribution margin per unit? How many units of that product should be produced each month?
   b. If Backyard could meet demand, what would be the total net operating income per month?
2. If demand was unlimited for all products, which products should Backyard Helpers produce?
3. Given the maximum number of machine minutes per month, use Solver in Excel to answer the following:
   a. How many of the following products should be produced each month?
      i. R25
      ii. JK369

b. Which products would be temporarily eliminated from Backyard Helpers product line under the Excel solution?
c. What is the total net operating income per month given the machine time constraints and the quantities determined by Solver?
d. What is the **maximum** amount Backyard Helpers should be willing to pay to rent fabrication time from another company? (Assume transportation and other costs would total $50,000.)

4. In Questions 3b, you identified products that would be temporarily eliminated if Backyard followed the Excel solution. What reasons, if any, might management have for continuing to produce some of those products even if it means reducing the supply of some of the other products?

LO3, 4

PB-24. Budget Variance Analysis Using Excel (Descriptive and Diagnostic Analytics)

Preston Township's City Council will be evaluating costs incurred in the various city departments at its next meeting. In total, costs exceeded budgeted amounts in the prior year by $576,277. The Council president has asked for information about actual vs. budgeted costs by department and by expense type to help in the evaluation process. Transaction and budget information is included in the **Budget Variance Data Set.xlsx** file available on the textbook's website. A video demonstrating Excel tools used to answer the questions in this problem is also available on the website.

1. Create two PivotTables.
   a. One for actual costs by department and expense type.
   b. One for budgeted costs by department and expense type.
2. Create two Budget Variance reports (one for variances by department and one for variances by expense type). Both reports should link actual and budget data from the PivotTables.
   a. The report should include columns for budgeted amounts, actual amounts, variance (in dollars), and variance (in percent). Show the unfavorable dollar variances as negative numbers, favorable variances as positive numbers. Show all percent variances as positive numbers. *HINT:* Use the ABS function in Excel in the formula to calculate percent variances.
3. Use the PivotTable and the variance reports to answer the following questions:
   a. Which department experienced the greatest variance between budgeted and actual cost (in dollars)? Which expense type in that department accounted for the largest share of the variance? *HINT:* Filter your PivotTables to update the variance reports.
   b. Which expense type had the highest unfavorable variance (in dollars)? Which department had the highest unfavorable variance in that expense type? Which expense type had the highest favorable variance (in dollars)? Which department had the highest favorable variance in that expense type?
   c. Which department had the highest percentage variance? Which type of expense was most over or under budget in that department? Which expense type had the highest percentage variance? Which department was most over or under budget in that expense type?
   d. Schools had the largest budget. Does it appear that the budget dollars were well managed? Explain your answer.
4. In general, should the council members be more concerned about the departments or expense types with the highest unfavorable dollar variances or the highest unfavorable percentage variances? Should the council members be concerned about departments or expense types with favorable variances? Explain your answers.

LO3

PB-25. Activity-Based Costing Using Excel (Predictive Analytics)

Kirkland Industries (a contract assembly manufacturer) has decided to adopt activity-based costing techniques to determine its manufacturing overhead rates. The production manager has identified three activities (materials movement, assembly, and packaging/shipping) and a number of possible activity measures (# of jobs, direct labor hours, machine hours, # of boxes shipped, and # of components used). Working together, the production and accounting managers have used historical data from 2014 to 2021 to determine total activity costs by month. Those results are included in the **ABC Cost Drivers Data Set.xlsx** file available on the textbook's website. The workbook also includes totals for the various activity measures from the same 2014-2021 period. (A video demonstrating Excel tools used to answer the questions in this problem is available on the website.)

Budgeted overhead dollars and activities are:

Budgeted Overhead		Budgeted Measures	
Materials movement	$1,080,000	# of jobs	480
Assembly	$1,950,000	Direct labor hours	16,000
Packaging/Shipping	$1,584,000	Machine hours	7,800
		# of boxes shipped	48,000
		# of components used	4,000,000

1. Use the correlation tool in Excel to determine which measure should be used for each activity. *HINT:* The correlation tool can be found on the *Analyze* menu on the Data tab in Excel. If the *Analyze* section does not appear, you will need to load the *Analysis ToolPak*. Click the *File* tab, click *Options*, and click *Add-Ins*. Make sure Excel *Add-ins* appears in the *Manage* field. Check the *Analysis ToolPak* option and click *OK*.
2. Using the measures identified in 1., determine the activity rates for allocating manufacturing overhead to jobs.
3. What would the predetermined rate be if direct labor hours were used to allocate all manufacturing overhead costs?
4. Assume Kirkland had a job that required 36 direct labor hours, 16 machine hours, 1,875 components, and 68 boxes. How much manufacturing overhead would be applied to that job under ABC? How does that compare to the amount applied if direct labor hours were used to allocate overhead? What might account for the difference?