C Data Analytics

Action Plan

LO	Topic/Subtopic	Page	Assignments
LO C-1	**Identify and define the four types of data analytics**	C-2	EC-1, EC-2, EC-3, EC-4, EC-5, EC-6, EC-7, EC-8, EC-9 EC-10, PC-11, PC-12, PC-13, PC-14, PC-15, PC-16, PC-17, PC-18, PC-19, PC-20, PC-21, PC-22, PC-23, PC-24
LO C-2	**Describe the use of data analytics within the accounting profession**	C-3	EC-1, EC-2, EC-3, EC-4, EC-5, EC-6, EC-7, EC-8, EC-9 EC-10, PC-11, PC-12, PC-13, PC-14, PC-15, PC-16, PC-17, PC-18, PC-19, PC-20, PC-21, PC-22, PC-23, PC-24

Data Analytics

LO C-1

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**. This technology can reveal trends and insights that would otherwise be lost in the overwhelming amount of data.

> **LO C-1 Overview**
>
> 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 C-1** summarizes the four types of data analytics.

Type of Data Analytics	Purpose	Example
Descriptive	To explain what happened	Sales trends for 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	Predicting fraudulent charges based on past behaviors
Prescriptive	To determine what should happen	Expand business to optimize sales based on predicted championship

EXHIBIT C-1
Type of Data Analytics

<table>
<tr><td>LO C-2</td><td>Data Analytics in the Accounting Profession</td></tr>
</table>

LO C-2 Overview

Describe the use of data analytics within the accounting profession.

Most accountants are likely already performing descriptive analytics in their regular work. This includes, for example, determining sum totals, averages, and period-to-period changes in such measures as net sales and net income. Other examples include average collection periods and day's sales in inventory by product type. Data analytics has other usefulness and should not be limited to only descriptive analysis. Accountants can provide far 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.

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. **Pricewaterhouse-Coopers** (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.

Summary of Learning Objectives

LO1 **Identify and define the four types of data analytics. (p. C-2)**
- 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. (p. C-3)   **LO2**
- Many accountants are already performing descriptive data analytics.
- Accountants can add much more 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.

Key Terms

Data analytics (p. C-2)   **Predictive analytics** (p. C-2)
Descriptive analytics (p. C-2)   **Prescriptive analytics** (p. C-2)
Diagnostic analytics (p. C-2)

Exercises

A video demonstrating the use of Microsoft Excel's built-in functions for data analysis is available on the textbook's website to assist in solving exercises EC-1 through EC-3.

Go to the book's website and download the file Employee data.xlsx, **link**. You will need to have the Analysis Toolpak add-in installed in Excel. It can be found under the Tools tab. If it does not appear, select Excel Add-ins under the Tools tab, and then check Analysis Toolpak. From the Excel ribbon, select Data and then Data Analysis. From the pop-up window, choose Descriptive Statistics, then click OK. Select the salary column as the input range, and check the box for labels in first row. Choose "New Worksheet" as the output option, click summary statistics, and click OK. Report the following:

a. Mean (average) salary
b. Median salary
c. Minimum salary
d. Maximum salary
e. Number of salary observations in the database

Exercise C-1
Using Microsoft Excel for Descriptive Analytics   **LO1, 2**

Go to the book's website and download the file Employee data.xlsx **link**. Place your cursor anywhere in the table of data. From the Excel ribbon, select Insert and then Pivot Table. The entire table should be selected automatically along with the choice to output the PivotTable to a new worksheet. Select OK. From the PivotTable Fields section, select and drag "Gender" and "Minority" to the "Rows" box below. Select and drag the variable "Education" to the "Columns" box. Select and drag the variable "Salary" to the "Values" box. Change the sum of salary to the average of salary by clicking the "i" icon to the right of the "Sum of Salary," choosing "Average," then clicking OK. Report the following:

a. Does additional education appear to be associated with a higher average salary?
b. Do males (1) or females (0) appear to earn higher average salaries?
c. Do minorities (1) or nonminorities (0) appear to earn higher average salaries? Does this hold for both genders?
d. What is the average salary of the entire population? What is the average salary of the entire population of males? What is the average salary of the entire population of females? What is the average salary of the entire population of male minorities? What is the average salary of the entire population of female minorities?

Exercise C-2
Using the Microsoft Excel PivotTable Function for Descriptive Analytics   **LO1, 2**

Go to the book's website and download the file Employee data.xlsx **link**. You will need to have the Analysis Toolpak add-in installed in Excel. It can be found under the Tools tab. If it does not appear, you will need to select Excel Add-ins under the Tools tab and then check Analysis Toolpak. From the Excel ribbon, select Data and then Data Analysis. From the pop-up window, choose "Regression," then click OK. Select values in the "Salary" column as the Input Y Range and values in the columns for "Gender" through "Education" for the Input X Range. Choose "New Worksheet" as the output option, then click OK. Report the following:

a. A measure on how well the independent variables gender, minority, and education are able to explain the variation in average salary is the adjusted R Squared. What percentage of the variation in average salaries is described by these variables?

Exercise C-3
Using Microsoft Excel for Diagnostic Analytics   **LO1, 2**

b. The t Stat is a measure of how an individual independent variable explains variation in the dependent variable average salary. An absolute value greater than 2 is generally considered a significant value in explaining variation. What do the t Stats tell us about the ability of the variables gender, minority, and education to explain average salary?

A video demonstrating the use of Tableau Public for data analysis is available on the textbook's website, to assist in solving exercises EC-4 through EC-6. In addition, a series of tutorial videos is available on the Tableau website https://public.tableau.com/en-us/s/resources.

Go to the Tableau website, https://public.tableau.com/en-us/s/, and download the app in order to complete exercises EC-4 through EC-7.

Exercise C-4
Using Tableau to Create Summary Statistics **LO1, 2**

Go to the book's website and download the file Employee data Tableau.xlsx, **link**. Connect Tableau Public to this Excel file. Go to the worksheet and drag the measures "Education," "Jobcat," and "Jobtime" up to Dimensions. Compare the average salaries by gender and minority status by dragging "Gender" to Rows and "Minority" to Columns and then "Salary" to the canvas. Change salary from a sum to an average.

a. How do average salaries compare by gender and minority status?

Next, explore how education level affects this relation by dragging "Education" to Columns. It may be easier to make this comparison by switching the order of "Minority" and "Education" on the Columns bar.

b. Does education level affect how average salaries compare by gender and minority status?

Next, change salary from average to maximum.

c. Does education level affect how maximum salaries compare by gender and minority status?

Exercise C-5
Using Tableau to Calculate a Visualization **LO1, 2**

Starting with the results from C-4, change salary back to average. Select the side-by-side bar chart (ninth selection) from the "show me" selections.

Based on this visualization of the data, what can you say about relative salaries for males and females and for caucasians and minorities?

Exercise C-6
Using Tableau to Forecast Future Values **LO1, 2**

Go to the book's website and download the file Employee data Tableau.xlsx, **link**. Connect Tableau Public to this Excel file. Go to the worksheet and drag the measures "Education," "Jobcat," and "Jobtime" up to Dimensions. Forecast average salaries for males and females based on Jobtime by first dragging "Gender" to Rows and "Jobtime" and "Salary" to Columns. Drag "Salary" to the canvas. Change salary from a sum to an average. From Analytics, select Forecast.

At 100 months of job time, how do the forecasts of average salary for males compare with females?

Exercise C-7
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.

Exercise C-8
Public Accounting Firms and Data Analytics **LO1, 2**

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

Exercise C-9
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.

Exercise C-10
Public Accounting Firms and Data Analytics **LO1, 2**

Go to EY.com and click on the three horizonal lines in the upper right corner. Next, choose "What we do" and then select "Our services" from the menu on the right side. Finally, click "Analytics and big data" from the popular quick links. Choose a topic and write about how Ernst & Young is using data analytics to help its clients.

Problems

Problems 11 through 24 utilize the widely used data analytics visualization software Tableau in order to answer a series of questions relating to various accounting and compensation topics. For those that do not have prior experience with the Tableau software, or those that wish to refresh their knowledge, the following videos are recommended. In addition to these recommended videos, Tableau offers many more free training videos on its website under the learning tab. These videos provide very useful information

about Tableau for those that wish to have a deeper understanding of this popular data analytics visualization software.

A general introduction to the software. (25 minutes). https://www.tableau.com/learn/tutorials/on-demand/getting-started?playlist=484034

An introduction to the Tableau interface. (4 minutes). https://www.tableau.com/learn/tutorials/on-demand/tableau-interface?playlist=484034

Gaining an understanding of relationships in order to connect to outside data. Stop at 14 minutes and 33 seconds. https://www.tableau.com/learn/tutorials/on-demand/relationships?playlist=484036

A general introduction to visual analytics. (6 minutes). https://www.tableau.com/learn/tutorials/on-demand/getting-started-visual-analytics?playlist=484037

How to use sorting. (5 minutes). https://www.tableau.com/learn/tutorials/on-demand/sorting?playlist=484037

An introduction to filtering. (2 minutes). https://www.tableau.com/learn/tutorials/on-demand/ways-filter?playlist=484037

A deeper look at filtering. (7 minutes). https://www.tableau.com/learn/tutorials/on-demand/using-filter-shelf?playlist=484037

Using interactive filters. (4 minutes). https://www.tableau.com/learn/tutorials/on-demand/interactive-filters?playlist=484037

An introduction to formatting. (7 minutes). https://www.tableau.com/learn/tutorials/on-demand/formatting?playlist=484037

The formatting pane. (7 minutes). https://www.tableau.com/learn/tutorials/on-demand/formatting-pane?playlist=484037

An introduction to calculation in Tableau. (3 minutes). https://www.tableau.com/learn/tutorials/on-demand/getting-started-calculations?playlist=484040

Calculation syntax in Tableau. (4 minutes). https://www.tableau.com/learn/tutorials/on-demand/calculation-syntax?playlist=484040

Problems 11–13 use data on executive compensation from S&P 500 companies for the years 2015 through 2019. The data is contained in the Excel file Compensation_Data_2015_2019_SP500 2.xlsx. The file can be downloaded from the book's website.

As a researcher in executive compensation you desire to learn more about the compensation differences between men and women in the roles of Chief Executive Officer (CEO) and Chief Financial Officer (CFO). Data including salary and total compensation, including salary, bonus, stock option awards, and miscellaneous income, for both men and women serving as CEOs and CFOs for S&P 500 companies for the five years 2015 through 2019 is contained in the Excel file Compensation_Data_2015_2019_SP500.xlsx. First connect to this file with Tableau and change the Year field from a number to a date type. Next create the following visualizations:

Problem C-11
Executive Compensation Visualizations with Tableau—Part I
LO 1, 2

a. A crosstab showing median salary by gender for the entire database. One method to accomplish this is to drag the dimension Gender to the rows shelf and then drag the measure Salary directly to the canvas. Next change the measure of Salary from Sum to Median. Do men or women have a higher median salary and by how much?

b. Create a crosstab that separates CEO median total compensation (Measure Comp1) by gender and sector. Also show how many individual CEOs are shown by gender and sector. One method to accomplish this is to drag the dimension Sector Name to the columns shelf and the dimensions Position and Gender to the rows shelf. Next drag the measures Comp1 and WRDS(Count) directly to the canvas. Finally change the measure of Comp1 to median. In the consumer Staples sector, how many men and how many women are CEOs, and what is their median total compensation? Note that the totals shown are combined for all five years.

c. Create a separate visualization in the form of a vertical bar chart that displays CEO total compensation by sector, position, and gender. Provide interactive filters for Year, Position, Gender, and Sector Name. Finally, use colors to highlight gender. Within the information technology sector, how much higher is the median compensation for male CEOs than for female CEOs in 2019? What is the name of the highest paid female CEO in this segment and where does she work? One method to accomplish this is to drag the dimensions Sector Name, Position, and Gender to the columns shelf. Next drag the measure Comp1 to the rows shelf

and change the measure to Median. Drag Year, Position, Gender, and Sector Name to the Filters shelf and then select Show Filter from each dimension pull down arrow. Select CEO from the Position filter and 2018 from the Year filter. Finally drag gender to the Color card and WRDS(Count) and Comp1 to the Label card. To see the data behind each of the bars simply hover over the bar and right click, then select Full Data.

d. Save the file for future use.

Problem C-12
Executive
Compensation
Visualizations with
Tableau—Part II
LO 1, 2

To further your research in executive compensation you would like to see the location by state, the level of executive compensation, and where females hold the position of CFO.

a. Create a map visualization that shows the location of companies that employ female CFOs. In 2016, which state had the highest level of median compensation for a female CFO? Who was the CFO, how much was she paid, and what company did she work for? One way to accomplish this is to first create a map by holding the control key on a Windows machine or the command key on a Mac and selecting the measures Latitude and Longitude, then selecting the map from the Show Me menu. Next drag Total Comp 1 over the Color card and change the measure to Median. Drag the dimension State over the Detail card and the measure WRDS(Count) over the Label card. Drag Gender, Position, Sector Name, and Year to the Filter shelf and select to show these filters. Select 2016 from the Year filter, CFO from the Position filter, and Female from the Gender filter. The state of Washington should appear the darkest indicating the highest median compensation. Click on this state and then click on the View data icon in the top right corner. Select the Full Data tab.

b. Save the file for future use.

Problem C-13
Executive
Compensation
Visualizations with
Tableau—Part III
LO 1, 2

In order to present your findings, you decide to construct a dashboard in Tableau that integrates the visualizations you created in Problems 11 and 12. You wish to allow the user of the dashboard to interact by having the bar chart serve as a filter for the map visualization.

a. Create a dashboard with the bar chart visualization created in Problem C-11 on the top and the map visualization created in Problem C-12 on the bottom. Make the bar chart serve as a filter for the map. What state had the most female CFOs in 2018 from the financials sector and what companies did they work for? How many male CFOs worked in the financials sector in 2018 in that state? Who was the only female CEO in the real estate sector in 2017 and in what state did she work and for what company? What was her total compensation in 2018? How many male CEOs were in the real estate sector in 2018 and what state employed the largest number? One way to create this dashboard is to open a new dashboard and drag in the bar chart visualization to the top and the map visualization under it. Next click on the bar chart visualization and click on the funnel icon to turn it on as a filter. In order to see the number of individuals in each bar you can go back to the bar chart and modify the visualization by dragging WRDS(Count) over the Tooltip card.

b. Save the file for future use.

Problems 14 through 16 use financial statement data for S&P 500 companies for the years 2015 through 2019. The data is contained in the Excel file Compustat_SP500_2015_2019.xlsx. The file can be downloaded from the book's website.

Problem C-14
Building Basic Tableau
Financial Accounting
Visualizations
LO 1, 2

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.

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.

Two widely used ratios used to analyze company performance are gross profit percentage and return on sales. You would like to create a visualization that compares these to ratios by segment and further compares segment performance to the median values of these ratios for the entire database of companies.

    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? 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? What company had the highest gross profit percentage in 2018 for the segment with the highest median gross profit percentage? One way to accomplish this is to start by going to the Analysis tab on the menu and selecting Create Calculated Field for both Gross Profit Percentage and Return on Sales. Next drag the dimension Segment to the columns shelf and the newly created measures to the rows shelf. From their pill pull down arrow change the Measure to Median. To allow the ability to filter on Year and Segment drag each of these dimensions to the Filter Shelf. Finally, to add overall company medians for each ratio switch from the Data pane to the Analytics pane and drag Median with 95% Cl to the canvas and place the distribution band on the pane option. Save the file for future use.

Problem C-15
Tableau Visualizations to Analyze Accounting Performance Measures
LO 1, 2

You have learned 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.

    Has the ratio days sales in inventory improved or declined in the consumer discretionary and the consumer staples segments between 2017 and 2018 and by how much? One way to determine this is to build two separate visualizations and then combine them in a single dashboard. The first visualization will show the level of the ratio for each segment for the two years in question. One way to create this first visualization is to first create the ratio Days sales in inventory. Next drag the dimensions Segment and Year to the columns shelf and drag the newly created ratio to the rows shelf and change its measure to median. In order to filter to be able to select specific segments and years, drag the dimensions Year and Segment to the Filters shelf and select the years 2017 and 2018 and the segments consumer discretionary and consumer staples. The second visualization, which will show the change in the ratio between 2017 and 2018, can be accomplished by first dragging the dimensions Segment and Year to the Columns shelf and also dragging the dimensions to the filters shelf and selecting the years 2017 and 2018 and the segments consumer discretionary and consumer staples. Next drag the ratio to the rows shelf and change its measure to median. In order to show the change between years use the pull-down arrow on the ratio's pill and select Quick table calculation and then Difference. Finally, right click on the year 2017 on the horizontal axis and select hide. As the last step, create a new dashboard and drag in the two visualizations. Save the file for future use.

Problem C-16
Using Tableau to Analyze Inventory
LO 1, 2

Hugo Enterprises has been performing its aging of accounts receivable manually; however, the task is becoming too time consuming. The company has recently acquired Tableau for some data visualizations; however, it was mentioned that the software could be used for receivables aging. You have been asked to perform Hugo's accounts receivable aging using Tableau. The file Aging Tableau 1.xlsx can be found at this link with Hugo's accounts receivable amounts and due dates.

Problem C-17
Using Tableau to Analyze Accounts Receivable   **LO 1, 2**

a. What is the total dollar value of the invoices that are between 31 and 60 days past due?

b. What is the largest invoice within the 91- to 120-day grouping?
    One way to accomplish this is to first link the Excel file to Tableau. By selecting the checkbox for using the built-in Data Interpreter the file can be easily cleaned up for further use. Now there should appear three columns with customer, amount, and due date data. Once this is done click on Sheet 1 (this can be renamed) in

order to begin creating the aging table. A calculated field will need to be created that computes the number of days each invoice is past due. Select Analysis and then select Create Calculated Field. Name this measure Past Due. Next use the formula DateDiff('day',[Due Date],#2023-12-31#) and click OK. Right-click on the newly created Past Due pill from the Measures shelf and then select Create and then select Bins. Change the size of the bins to 30 and click OK. Now double click the Past Due (bin) in the Dimensions shelf to see a list of the aging groups. To fill in the table with invoice amounts simply double-click on the Amount measure in the Measures shelf. In order to see the data that makes up each total in the aging simply click on the amount and then select the view data icon in the top right corner and select Full Data at the bottom of the pop-up window. In order to see a visualization of the data simply select an appropriate chart from the show me selections. Save the file for future use.

Problem C-18
Using Tableau to Detect
Fraud with Benford's
Law   **LO 1, 2**

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 dataset 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.

Part A. Download the file GDP Tableau.xlxs 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. In order to use Benford's Law you first need to connect this file to Tableau. Once connected go to the worksheet Sheet 1 (rename if you wish) where you will need to create a couple of calculated fields. The first 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). 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 next drag First Integer from the Dimensions area to Columns and drag Number of Records (Table 1 (Count)) from the Measures area to Rows. Click CNT(Table 1) on Rows to show the pulldown 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 popup. 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 names 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 to Benford's Law?

Part B. 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.xlxs from the textbook website and follow the same procedure as you did in part A above.

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

Part C. 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.xlxs 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.

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

   e.   Save the file for future use.

Problems 19 through 24 use financial statement data for S&P 500 companies for the years 2015 through 2019. The data is contained in the Excel file Compustat_SP500_2015_2019.xlsx. The file can be downloaded from the book's website.

One of the ratios that provides information on how well a company utilizes its assets is the ratio asset turnover. As part of your analysis of different segments of the S&P 500 you would like to have a visualization that ranks companies within various segments by their asset turnover.

Problem C-19
Using Tableau to Analyze Fixed Assets
LO 1, 2

   a.   For the year 2019 what firm in the consumer discretionary segment had the second highest asset turnover? How did the same firm rank in 2018? Did the segment consumer discretionary or consumer staples have a higher median level of asset turnover in 2017? One way to determine this is to create a visualization by first creating a calculated field of the ratio asset turnover and then dragging the dimensions Segment and Company Name to the rows shelf and the ratio asset turnover to the columns shelf. Using the pull-down arrow of the ratio's pill, change the measure to median. Next sort the horizontal bar chart by Company Name. Drag the dimension Year to the filter shelf and select Show filter. Finally switch from the Data pane to the Analytics pane and drag Median with 95% Cl on the canvas and place the distribution band on Pane. Save the file for future use.

A popular ratio for analyzing a company's short-term liquidity is the current ratio. In your continuing analysis of S&P 500 companies you have decided to build a simple visualization that uses color highlighting and tooltip labeling to show current ratios by segment over the five-year period 2015 through 2019. You also desire to do some further analysis in Excel, so you will want to export this data as a crosstab to Excel. Finally, you want to include your visualization in a PowerPoint presentation.

Problem C-20
Using Tableau to Analyze Liquidity
LO 1, 2

   a.   What segment has the highest current ratio in each year? One way to create this visualization is to first create the calculated field current ratio. Next drag the dimension Year to the Columns shelf and the ratio Current ratio to the Rows shelf. Next change the ratio's measure to median. Finally drag the dimension Segment over the Color card. In order to export the data as a crosstab to Excel click on the worksheet tab in the menu bar and choose export > crosstab to Excel. If you would prefer to see the data as a crosstab in Tableau, right click on the Tableau workbook icon and choose Duplicate as crosstab. Finally, to export the visualization to PowerPoint, select File from the menu bar and choose Export as PowerPoint. Save the file for future use.

Many equity investors are particularly interested in the dividend paying policy of a company. Two ratios that provide information in this regard are the dividend payout ratio and the dividend yield. As an analyst you would like to build a visualization that looks at these two ratios together and in particular which segments outperform the S&P500 in general.

Problem C-21
Using Tableau to Analyze Dividend Policy **LO 1, 2**

   a.   For the year 2015, did any segment outperform or underperform the median S&P500 company for both the dividend payout and dividend yield by more than a 95 percent confidence level? One way to determine this is to first create calculated fields for both dividend yield and dividend payout and then build a scatter graph visualization with median dividend payout on the columns shelf and median dividend yield on the rows shelf. Next drag Segment to the Color card to color highlight the segments. Drag Year to the filter shelf in order to filter on Year. Finally, switch to the Analytics pane and drag Median with 95% Cl to the canvas. Save the file for future use.

One of the basic tools in any analyst's toolkit is the DuPont method of Return on Equity (ROE) decomposition. You would like to build a series of visualizations to exploit the DuPont method to find value in equities. To do this you will need visualizations that show, by segment, both ROE and the components of ROE, namely Return on Sales (ROS), Asset Turnover (AT) and Leverage (LEV).

a. Which segment had the highest ROE in both 2017 and 2018? Which component of ROE was mostly responsible? The first step in determining this is to create calculated fields for ROE, ROS, AT, and LEV. The next step is to create a vertical bar chart visualization by dragging the dimensions Segment and Year to the Columns shelf and the median measure of the four new calculated fields to the rows shelf. The dimension Year needs to be dragged to the filter shelf with the years 2017 and 2018 selected. Finally, Segment can be dragged over the Color card to highlight by color.

b. Which segment showed the largest gain and the largest decline in ROE between 2017 and 2018? This is best shown on a dashboard by combining two separate visualizations. The first visualization in the form of a vertical bar chart can be constructed by dragging the dimensions Segment and Year to the Column shelf and the measure of Median ROE to the rows shelf. The dimension Year should be dragged to the filters shelf and the years 2017 and 2018 selected. To make segments easier to see in the visualization, Segment can be dragged over the Color card to highlight by color. The second visualization, which will show the change in the ratio between 2017 and 2018, can be accomplished by first dragging the dimensions Segment and Year to the Columns shelf and also dragging the dimension Year to the filters shelf and selecting the years 2017 and 2018. Next drag the ratio ROE to the rows shelf and change its measure to median. In order to show the change between years use the pull-down arrow on the ratio's pill and select Quick table calculation and then Difference. Finally, right click on the year 2017 on the horizontal axis and select hide. To make segments easier to see in the visualization, Segment can be dragged over the Color card to highlight by color. As the last step, create a new dashboard and drag in the two visualizations.

c. Save the file for future use.

You wish to continue your analysis of segment ROE by looking at trends within segments, then identifying what companies rank the highest for ROE within those segments.

a. For the period 2015 through 2019, which segments showed the best positive trends in ROE growth? One way to create a line chart visualization showing this trend is to drag the dimensions Segment and Year to the Columns shelf and drag the Median measure of ROE to the Rows shelf. To make segments easier to see in the visualization, Segment can be dragged over the Color card to highlight by color. Finally, switch to the Analytics pane and drag Trend line to the canvas, selecting "linear."

b. For the segments identified in part a, which firms reported the highest ROE in 2019? A horizontal bar chart visualization can be created by dragging the dimensions Segment and Company name to the Rows shelf and the Median measure of ROE to the columns shelf. Segment and Year can be dragged to the filters shelf to allow filtering on these two dimensions. Finally, to aid in the analysis, the company names can be sorted by clicking on the sort icon located on the tool bar.

c. Save the file for future use.

To complete your analysis of the DuPont method you would like to have an interactive dashboard that allows the user to select any of the S&P 500 companies and see its segment, its change in ROE between 2017 and 2018, along with the changes in each of the items making up components of ROE.

Problem C-24
Using Tableau to Study ROE with the DuPont Method—Part III
LO 1, 2

a. How much did Alphabet's (the parent company of Google) ROE improve between 2017 and 2018 and which component showed the largest increase? Pick any other company and answer the same question. One approach to build this dashboard is to first create five visualizations. The first horizontal bar chart visualization can be created by dragging the dimensions Company name and Year to the Rows shelf and the measure ROE to the Columns shelf. The dimension Year needs to be dragged to the filters shelf and the years 2017 and 2018 selected. Next drag the dimension Company name to the filters shelf and click on Show filter from the pills pull-down menu. To allow segments to be highlighted by color in the visualization, Segment can be dragged over the Color card. This process can be repeated to create separate visualizations for the components of ROE, namely ROS, AT, and LEV. The fifth visualization, a line chart, can be created by dragging the dimension Year to the Columns shelf and the measures Net Income, Stockholders Equity, Sales, and Total Assets to the rows shelf. The dimension Year should be dragged to the filters shelf and the years 2017 and 2018 selected. The dimension Company name should also be dragged to the filters shelf and then have Show filter selected. Finally, to allow segments to be highlighted by color in the visualization, Segment can be dragged over the Color card. Now that each of the visualizations are complete, create a dashboard by dragging the first four visualizations, one under the other, on the left side of the dashboard and drag the fifth visualization to the right side of the dashboard. Convert the filter of Company name to a Single Value drop-down floating filter by selecting those two options and position the filter at the top of the dashboard. Save the file for future use.