CHAPTER

8

HEALTHCARE ANALYTICS

Operations Management in Action

Analytics has played a pivotal role in combatting the spread and impact of COVID-19. Two examples follow. First, the Centers for Disease Control and Prevention (CDC) created a live dashboard of COVID-19 statistics to track the number of cases, hospitalization rates, and mortality rates broken down by state, county, gender, race, and age. These dashboards were pivotal in not only showing the number of infections but also in managing disease hot spots. When cases rose rapidly in certain areas of the county, the CDC could adjust its guidelines and inform local authorities of the greater transmission rates. The CDC developed a separate analytics dashboard to show the number and percentage of people who received a COVID 19 vaccination (exhibit 8.1).

Second, predictive analytics models are being used to identify high-risk COVID-19 patients in order to reduce the severity of their cases. One of the unexpected adverse outcomes of COVID-19 was the stop in elective procedures as resources were redirected to counter the pandemic. Health Catalyst indicates that hospitals lost more than $60 billion monthly because of this halt, which also has had an enormous effect on patient health (Health Catalyst Editors 2021). Using machine learning and other predictive models, Health Catalyst has been able to quantify the impact of primary care provider disruptions on health and financial outcomes. These models help to prioritize high-risk patients to ensure they are getting the proper care and reduce the overall cost of their healthcare. Exhibit 8.2 is a simple predictive-model output that clearly shows the first patient had a primary care provider disruption that could incur significant costs. Such patients then can be prioritized for care to help reduce the mortality rate and financial impact of COVID-19 on the health system.

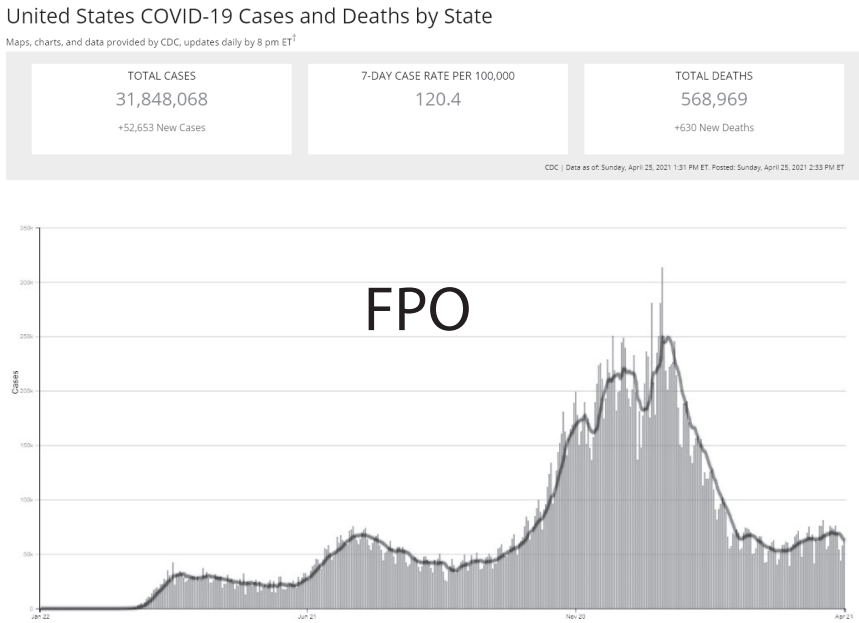
OVERVIEW

"Too much data and not enough information" has never resonated more than in today's healthcare environment. In response, the disciplines of analytics, big data, and informatics have exploded and even become commonplace in hospital and health system operations.

What Is Analytics in Healthcare?

In 2007, Thomas Davenport and Jeanne Harris wrote their seminal book, *Competing on Analytics: The New Science of Winning*. This text demonstrates

EXHIBIT 8.1
United States COVID-19 Cases and Deaths by State



Source: Centers for Disease Control and Prevention (2021).

how companies from many different industries can use analytics to create value and improve organizational performance.

Analytics, defined by one source as the "the systematic computational analysis of data or statistics" (Oxford Living Dictionaries 2016), has become particularly popular across the healthcare landscape for a number of reasons:

- More data than ever before are generated and available—particularly with the wide adoption of electronic health records (EHRs).
- The current regulatory environment requires the reporting of thousands of measures.

EXHIBIT 8.2
A Predictive Model with PCP Disruption as a Feature

Age	Gender	Diabetes	CHF	Had PCP Visit Disruption		Total Cost (Target Outcome)
79	M	Y	N	Y	➡	$100,000
79	M	Y	N	N	➡	$20,000
65	F	N	Y	Y	➡	$30,000

Source: Exhibit from "Data Science Reveals Patients at Risk for Adverse Outcomes Due to COVID-19 Care Disruptions" (on https://www.healthcatalyst.com/insights/healthcare-data-science-reveals-high-risk-patients) by Health Catalyst, Inc., and is used under license. Note: CHF = congestive heart failure; PCP = primary care provider.

- Hospitals and health systems are facing increased pressure to improve clinical, operational, and financial results.
- Population health has become a competitive strategy, and analytics is crucial to shaping effective population health initiatives.
- Information technology and software are increasingly sophisticated, allowing analysis of data on a massive scale.

More Data

Due to ever-increasing computing power and the advent of cloud storage, smartphones, and other technologies, more data and information are available today than ever before. This availability presents both challenges and opportunities in data storage, security, and management. In large part because of the availability of funds—and new mandates—from the American Recovery and Reinvestment Act of 2009, most hospitals and clinics have installed EHR systems. The massive conversion from paper charts and records was difficult for many organizations to accomplish, but EHRs are finally stable enough to be used as a good data resource. Epic Systems Corporation and Cerner Corporation are the two largest software companies to have created platforms to store health records. The widespread use of these systems has given healthcare providers the capability to longitudinally collect data on patients, which offer healthcare systems comprehensive insights and, potentially, the capability to improve care decision making.

Regulatory Environment

The effective use of analytics can help healthcare organizations manage mounting regulatory pressures. The Centers for Medicare & Medicaid Services require every hospital to report approximately 1,700 quality measures for regulatory compliance (Blumenthal, Malphrus, and McGinnis 2015). The sheer number of data points that must be collected forces organizations to dedicate significant resources to collecting and managing the data. And this effort does not take into account the additional resources required to analyze and make decisions with the data.

Pressure to Produce Results

In Minnesota, a unique relationship was formed between Allina Health and Health Catalyst. Health Catalyst provides analytics services to Allina to assist in project management, continuous improvement, population health analysis, and financial analytics. The use of large-scale data allows the healthcare system to focus on achieving results through coordinated efforts. Data have been used to analyze a variety of system elements, including doctors' efficiency, clinic efficiency, and overall system effectiveness.

Healthcare regulation and competition pressure have changed the marketplace for health systems. Organizations need a systematic approach to

reducing costs and finding new market opportunities. The operations improvement tools discussed in many of the chapters of this book can be made more powerful with the use of advanced analytics.

Population Health

Kindig and Stoddart (2003) define population health as "the health outcomes of a group of individuals, including the distribution of such outcomes within the group" (p. 380). The increased use of EHRs gives health systems the ability to understand the costs and clinical trends related to the patients they serve. This capability allows the development of specific treatments for diseases and conditions, which leads to improved outcomes. One of the hallmarks of big data analysis in healthcare is the use of predictive models.

Winters-Miner (2014) identifies seven ways predictive analytics can improve healthcare:

- Improves diagnosis
- Helps with preventive medicine and public health efforts
- Provides answers to physicians for the treatment of individual patients
- Provides employers and hospitals tools to predict insurance product costs
- Allows smaller test cases to be used to prove models
- Helps pharmaceutical companies meet the needs of the public for medication
- Potentially helps improve outcomes

Sophisticated Technology

Technology breakthroughs are enabling analysts to tackle increasingly complex problems. Analytics technology not only allows larger data sets to be used but also increases the speed at which analysis can be completed. The advanced technology used in analytics today can enable organizations to perform better and more sophisticated analytics. Data visualization software can easily replicate dashboard charts and graphs with new data. Statistical software can find relationships in large data sets once too difficult to analyze. Such software is a critical tool in the development and deployment of advanced analytics in healthcare.

Introduction to Data Analytics

The goal of data (big and small) analytics is to obtain actionable insights that result in smarter decisions and better business outcomes. Many of the tools and techniques from chapters 7 and 9 can be applied in an analytics environment.

The basic work of an analyst is to build a data framework through the following major goals:

- *Gathering data*—Data are the facts provided by databases.
- *Building information*—Information is the layer on top of data that helps make sense of the data. Without essential knowledge of the business situation, the information is likely not valuable.
- *Gaining actionable insights*—Actionable insights are those nuggets of knowledge from the information that affect the organization. The insights should enhance a leader's ability to make improved decisions.

This framework for data analysis provides the background for the various forms of analytics.

What Is Statistical Thinking?

As defined by Joseph Juran, statistical thinking is the collection, organization, analysis, interpretation, and presentation of data (Juran and De Feo 2010). In most business systems in the healthcare industry, statistical thinking is lacking. Knowledge-based management and improvement require that decisions be based on facts rather than on feelings or intuition. Collecting the right data and analyzing them correctly enable fact-based decision making.

The importance of understanding statistical concepts in developing high-performing healthcare systems cannot be overstated. Delivering high-quality healthcare in a sustained manner depends on understanding and controlling variance. Variance is present in all systems, but the ability of leadership to understand and control variance distinguishes high-performing systems from poorly run systems. *The irony of this relationship is that many clinical quality and safety rules and regulations are designed and driven by the understanding of variance, whereas the supporting business systems are often designed simply to meet regulatory agency requirements and not to manage the variance in the system. The good news is that this situation provides the opportunity to make massive changes to system and financial performance simply by understanding data and metrics.*

Analytics can be described as taking place in three distinct phases:

- Descriptive analytics
- Predictive analytics
- Prescriptive analytics

Descriptive Analytics

Descriptive analytics is the process of condensing large data sets into meaningful information that can assist in decision making. Descriptive statistics examine past performance and summarize data to discern trends and patterns to explain

behavior. In healthcare, reporting mechanisms such as regulatory compliance, quality measures, and financial results commonly use descriptive analytics.

Descriptive analytics makes up the largest subset of the analytics field. One main feature of data visualization is making data consumable by people. The process of converting raw data is necessary because data alone are not typically usable to managers.

Examples of descriptive analytics outputs include the following:

Business intelligence The process of converting raw data through a variety of methods into information that can assist with decision making.

- Business intelligence reports
- Dashboards with key performance indicators (KPIs)
- Descriptive statistics
- Traditional data visualization techniques

Predictive Analytics

Predictive analytics builds models on the basis of data that can help forecast the future in terms of probabilities. Models cannot perfectly predict the future but can provide insights for individuals to make effective decisions. Predictive analytics uses a variety of statistical techniques, ranging from regression modeling to machine learning to data mining, to make projections about future events.

In healthcare, the use of predictive models has become popular in disease management and population health. For example, some healthcare organizations have begun to examine early indicators of diabetes to help prevent and lower costs associated with diabetes management (Barton 2021). This analytics activity is important because, according to the CDC (2009), more than 75 percent of total healthcare spending in the United States is related to chronic healthcare conditions.

At Hennepin County Medical Center (HCMC), population health analysts discovered that individuals diagnosed with HIV also suffered from poor nutrition. A predictive model was constructed that showed the positive impact of improved nutrition on healthcare costs. Today, HCMC distributes healthy food with HIV medications to many of the patients in this population and have found total costs to be reduced.

In short, predictive models have become a common approach to help reduce costs, improve quality outcomes, and lower overall patient risk.

Predictive Tools

Three approaches are typically used for developing predictive models: regressions, decision trees, and neural networks.

Regressions

A number of regression-type approaches can be used to predict future performance from historical data. Most analytical software (e.g., SAS, SPSS) packages include numerous regression tools.

Decision Trees

Decision trees are a form of "supervised learning" tools. The decision tree algorithm first suggests a split of the databases into a series of "leaves," whereby each data point is allocated to one leaf. If the analyst agrees with the computer's selection of leaves, the computer then suggests a further subdivision of the leaves. This process continues until the analyst believes the full tree represents a good model of the data.

Although regressions may be more accurate in their predictive capability, decision trees are useful for explaining the predictions to nonanalysts. A version of the decision tree tool was used to create the Medicare diagnosis-related group (DRG) system in 1983. Exhibit 8.3 demonstrates the use of a decision tree to predict annual costs for Medicare patients.

Neural Networks

Neural networks attempt to mimic the human brain in the following ways:

- Input units obtain the values of input variables and, if the analyst chooses, standardize those values.
- Hidden units perform internal computations, providing the nonlinearity that makes neural networks powerful.
- Output units compute predicted values and compare those predicted values with the values of the target variables.

Units pass information to other units through connections. Connections are directional and indicate the flow of computation in the network.

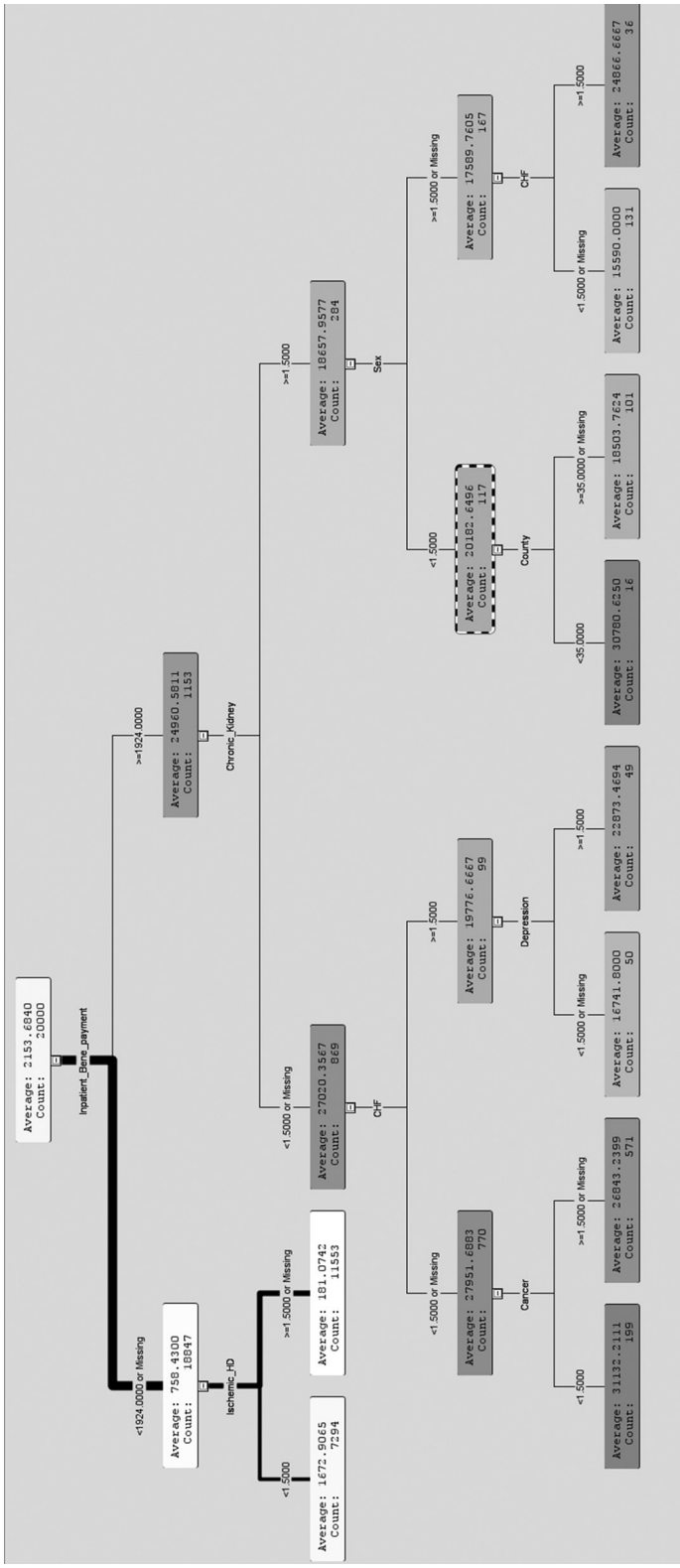
Once a neural network is created, it can be applied to predict outputs on the basis of new inputs. A challenge in using neural networks is that they are sensitive to the initial data used to calibrate the network. In addition, because of the hidden computations, neural networks are difficult to diagnose and correct if they are not operating properly.

Prescriptive Analytics

Prescriptive analytics provides decision makers with models that offer guidance in the form of recommendations. These models use a combination of predictive models, optimization, mathematical models, and other techniques to generate prescriptive solutions. Examples of prescriptive models include the following:

- Models for staffing that maximize quality outcomes and minimize costs
- Models to maximize capacity in operating rooms
- Strategic models that demonstrate efficient allocation of capital investments
- Risk models that minimize adverse health events

EXHIBIT 8.3
Decision Tree Example



Note: Diagram created with SAS Miner. Data are simulated Medicare diagnosis and annual inpatient payments.

Healthcare problems are complex and multidimensional and can be difficult to model. In the modeling process, many assumptions are made in prescriptive models such as optimization. Decision makers can use prescriptive models in combination with their knowledge of the healthcare system to make effective decisions.

Data Visualization

Data visualization tools help decision makers extract value from raw big data. They enable users to quickly view, and make sense of, large amounts of data and to combine several data sources.

When dealing with most real-world data sets, the analyst can expect to spend up to 80 percent of her time finding, acquiring, loading, cleaning, and transforming data. Some of this process can be performed with automated tools, but almost any data cleaning involving two or more data sets requires some level of manual work.

Many forms of data visualization have been developed. Those discussed in this section include traditional charts and graphs and dashboards. These examples represent just a few of the common forms of visualization used today in hospitals and health systems.

Traditional Charts and Graphs

Bar Graphs

Bar graphs, or column graphs, help users visualize the scale of differences between categories. Exhibit 8.4 is a bar graph showing how much a hospital

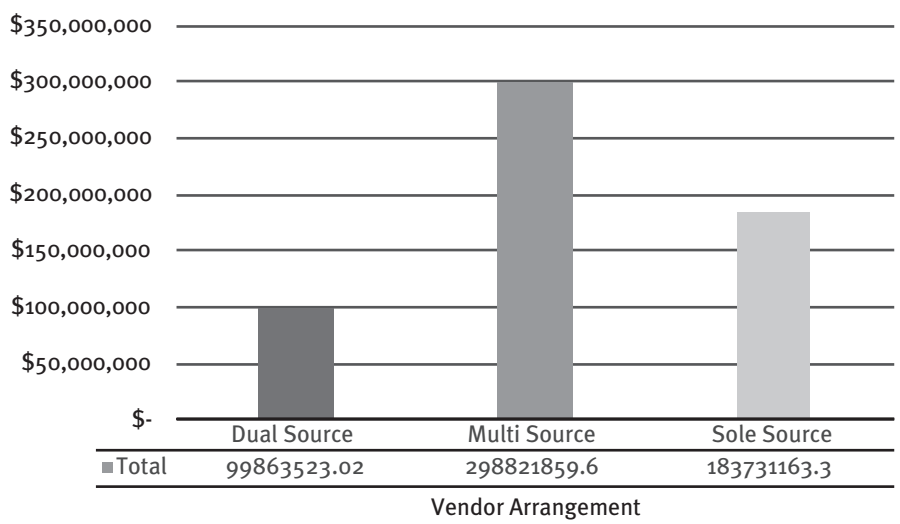
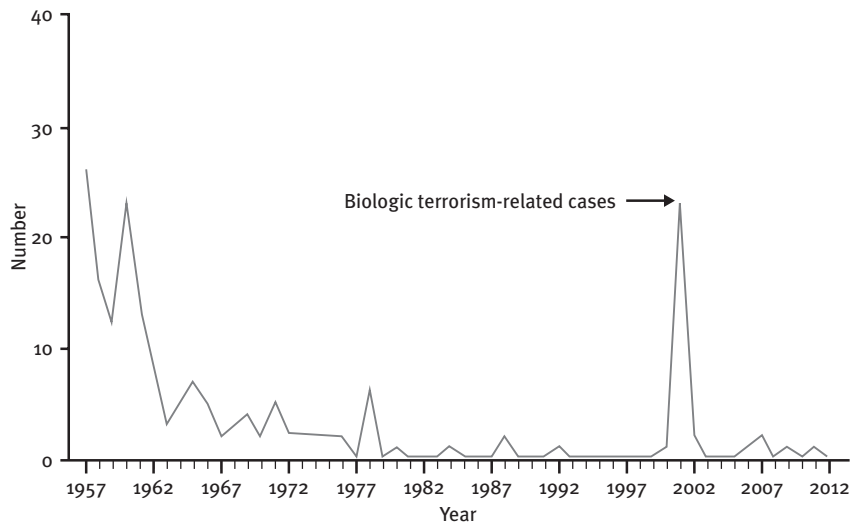


EXHIBIT 8.4 Bar Graph Showing Total Allocation by Vendor Type

EXHIBIT 8.5
Line Graph Showing Number of Biological Agent Cases Reported, 1957–2012



Source: Adams et al. (2014).

system is spending on purchasing by vendor type. This is a classic example of a traditional business intelligence report created in Microsoft Excel.

Line Graphs

Another traditional business intelligence report is a classic line graph. Line graphs are useful in examining data over time. Exhibit 8.5 is a line graph showing the number of cases of biological agents reported to the CDC from 1957 to 2012. The peak in the early 2000s represents the anthrax cases reported in the time frame following the 9/11 terrorist attacks on New York City and Washington, D.C., in 2001. As the exhibit demonstrates, line graphs reveal opportunities to explore trends and peaks in activity.

Map Functionality

Exhibit 8.6 is an example of a map of diabetes concentration by county in the United States created in Tableau. Tableau is powerful data analysis and visualization software that allows a user to create pictures by inputting data. Although such mapping does not have any predictive capability, exhibit 8.4 demonstrates its effectiveness in showing, for example, where the highest concentrations of reported individuals with diabetes reside. These types of maps help decision makers understand the concentration of data in geographic locations.

Histograms and Scatter Plots

Scatter plots show the relationships between two variables, and histograms are graphical representations of the distribution of data. These visualization techniques are covered in more detail in chapter 9.

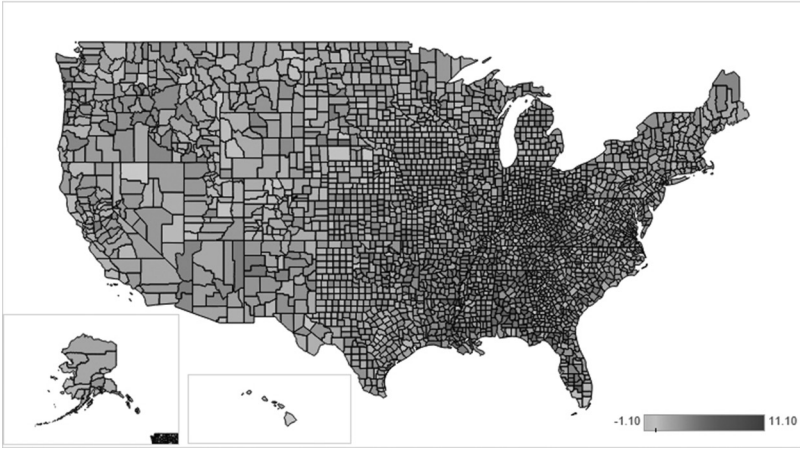


EXHIBIT 8.6
Interactive Map of Diabetes Prevalence by US County, 2004–2012

Source: Cook (2015). Used with permission.
Note: Higher numbers and darker grayscale indicate an increase in diabetes prevalence.

Dashboards[1]

A key purpose of an analytics department is to collect and define metrics and KPIs for executive and operational dashboards. Although the techniques discussed here can be used across many different business intelligence gathering efforts, they are also useful for collecting and organizing business data into a format for effective dashboard design.

With the explosion of dashboard tools and technologies in the business intelligence market, many people have different understandings of what a dashboard, metric, and KPI consist of. In an effort to create a common vocabulary, we define a set of terms that form the basis of our discussion. Although the definitions provided in the following subsection might seem onerous and require a second reading to fully understand them, once grasped, these concepts avail you of a powerful set of tools for creating dashboards with effective and meaningful metrics and KPIs.

Metrics and Key Performance Indicators

Metrics and KPIs are the building blocks of many dashboard visualizations, as these components are the most effective means of alerting users to their progress toward achieving their objectives. In addition to being the products of an organization's goals and objectives, metrics and KPIs may arise from strategy maps (discussed in chapter 5).

The definitions that follow build from one concept to the next and help inform dashboard design. Take the time to understand each definition and the related concepts before moving on to the next definition.

Metrics

The term *metric* refers to a direct numerical measure that represents a piece of business data in relationship with one or more dimensions. One example is gross sales by week. The measure is dollars (gross sales), and the dimension is time (week). For any given measure, viewing the values across different hierarchies in a dimension may be helpful. For instance, a display of gross sales by day, week, and month shows the dollars (gross sales) measure along different hierarchies (day, week, and month) in the time dimension. The term *grain* refers to the association of a measure with a specific hierarchical level in a dimension.

Looking at a measure across more than one dimension, such as gross sales by territory and time, is called multidimensional analysis. Most dashboards do not leverage multidimensional analysis except in a limited and static way; more dynamic "slice and dice" tools are available in the business intelligence market. This qualification is important to note. Say you uncover a significant need for this type of analysis in the requirements gathering process. Knowing that these robust tools exist, you have the option of supplementing your dashboards with some type of multidimensional analysis tool.

Key Performance Indicators

A KPI is simply a metric that is tied to a target. Most often, a KPI represents the distance a metric is above or below a predetermined target. KPIs usually are shown as a ratio of actual to target and are designed to instantly let a business user know if he is on or off track without having to consciously focus on the metrics represented. For instance, an organization may decide that, to hit the quarterly sales target, it needs to sell $10,000 worth of syringes per week. The metric is syringe sales per week, and the target is $10,000. Using a percentage gauge visualization to represent this KPI, and assuming we had sold $8,000 in syringes by Wednesday, the user would instantly see that he is at 80 percent of the goal.

When selecting targets for KPIs, remember that a target is needed for each grain you want to view in a metric. Having a dashboard that displays a KPI for gross sales by day, week, and month, for example, requires that targets be identified for each associated grain.

Scorecards, Dashboards, and Reports

The difference between a scorecard, a dashboard, and a report can be one of fine distinctions. Each of these tools can combine elements of the other, but at a high level they all target distinct and separate levels of the business decision-making process.

Scorecards

Starting at the highest, most strategic level of the business decision-making spectrum are scorecards. Scorecards are primarily used to help align operational

EXHIBIT 8.7
Sample Hospital Scorecard

Goal	Target	Owner	Review Frequency	Aug-16	Sep-16	Oct-16	YTD 2016
Finance							
Patient Information Accuracy Rate	99%	Paul	Monthly	99%	100%	97%	100%
Denials and Write-offs as % of Overall Charges	4%	Sarah	Monthly	5%	4%	3%	5%
Number of Days Charged in A/R	5	Sarah	Monthly	2	6	1	4
People							
Absenteeism Hours	30	Joseph	Monthly	15	20	30	22
Acceptable Overtime Hours	7%	Joseph	Monthly	8%	4%	5%	6%
Staffing: Open Positions	3	Jennifer	Monthly	2	1	1	1
Clinical							
Hospital-Wide 30 Day Readmissions	10.0%	Mark	Monthly	13.0%	11.0%	9.8%	12.2%
Heart Failure Mortality	13.2%	Mark	Monthly	12.7%	11.0%	9.0%	10.7%
Inpatient LOS (Days)	3	Catherine	Monthly	2.7%	2.3%	2.6%	2.5%

Source: Exhibit from "Healthcare Dashboards vs. Scorecards: Use Both to Improve Outcomes" (at www.healthcatalyst.com/healthcare-dashboards-vs-scorecards-to-improve-outcomes) by Health Catalyst, Inc., and used under license.

execution with business strategy. The goal of a scorecard is to keep the business focused on a common strategic plan by monitoring real-world execution and mapping the results of that execution back to a specific strategy (see chapter 5). The primary measurement used in a scorecard is the KPI. These indicators are often a composite of several metrics or other KPIs that measure the organization's ability to execute a strategic objective.

Exhibit 8.7 is an example of a hospital scorecard. In this example the scorecard demonstrates how well the hospital is doing on key strategic measures (KPIs). By examining this scorecard, a decision maker can quickly observe that the 30-day readmission rate needs improvement.

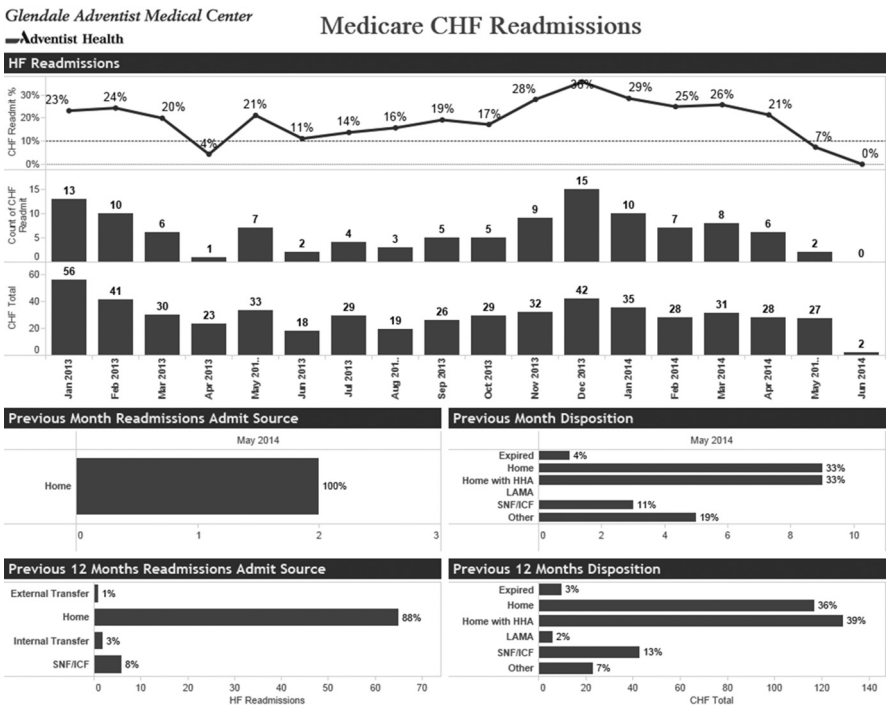
Dashboards

A dashboard resides one level down from a scorecard in the business decision-making process, as it is less focused on a strategic objective and more tied to operational goals. An operational goal may directly contribute to one or more high-level strategic objectives. In a dashboard, execution of the operational goal itself becomes the focus, not the high-level strategy. Dashboards are a key tool in the implementation of the balanced scorecard discussed in chapter 5.

The purpose of a dashboard is to provide the user with actionable business information in a format that is both intuitive and insightful. Dashboards leverage operational data primarily in the form of metrics and KPIs.

Exhibit 8.8 shows an example of a sample dashboard created using a popular data visualization software (Tableau). The decision maker can easily examine the visual and determine the status of readmissions for the organization. The key to the dashboard is that it presents real-time information that enables the organization to act quickly. What makes the dashboard different from the scorecard is both the granularity of the report and the individuals

EXHIBIT 8.8
Sample Tableau Dashboard



Source: Reprinted with permission from Corbett (2014).

for whom the report is intended. Scorecards are often used by organizational leaders, whereas dashboards are often used by operational managers who make daily decisions.

Reports

Probably the most prevalent business intelligence tool seen in business today is the traditional report. Reports can be simple and static in nature, such as a list of sales transactions for a given period, or more sophisticated cross-tab reports with nested groupings, rolling summaries, and dynamic drill-through or linking. Reports are most appropriate when the user needs to look at raw data in an easy-to-read format.

When combined with scorecards and dashboards, reports allow users to analyze the specific data underlying their metrics and KPIs.

Gathering Key Performance Indicator and Metric Requirements for a Dashboard

Traditional business intelligence projects often take a bottom-up approach in determining requirements, where the focus is on the domain of data and the relationships that exist in those data. When collecting metrics and KPIs for your dashboard project, however, taking a top-down approach is preferred.

A top-down approach starts with the business decisions that must be made first and then works down into the data needed to support those decisions. To take a top-down approach, you must involve the business users who will be using these dashboards, as these are the only people who can determine the relevancy of specific business data to their decision-making process.

Data Mining for Discovery

The vast majority of work being performed in healthcare analytics today is in reporting (descriptive analytics), with some highly specialized work in predictive and prescriptive analytics. Almost all of these tasks share a common characteristic: They entertain a specific hypothesis. Examples are as follows:

- I believe that patients of some doctors experience significantly longer lengths of stay than those of other doctors for the same DRG.
- I believe I can predict the amount of time a health plan will take to remit payment.
- I believe I can predict which patients will not fill their prescriptions on the basis of their zip code.

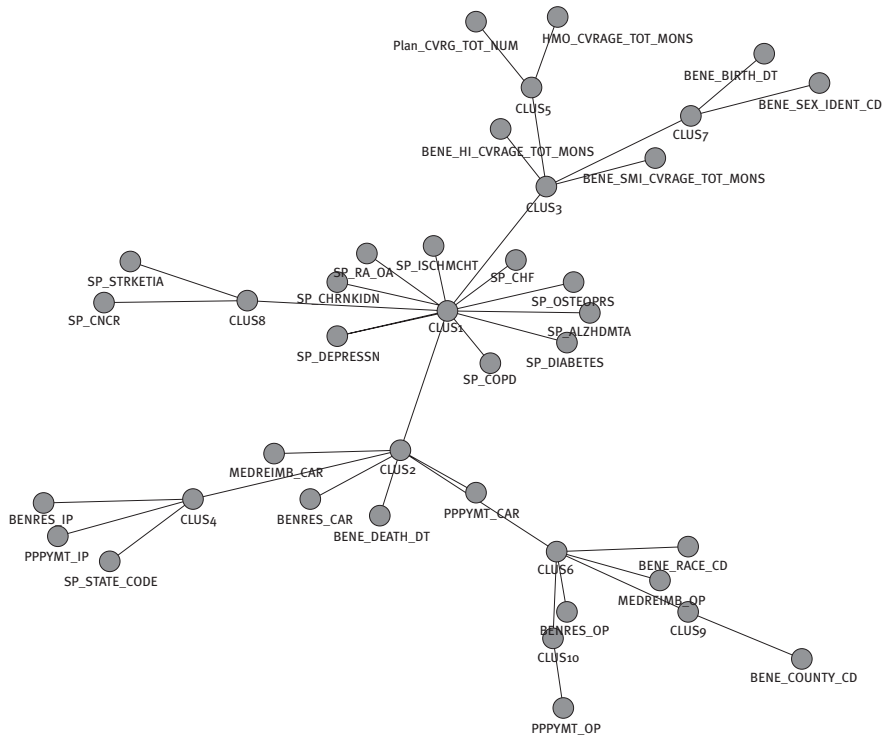
However, another powerful approach—data mining—is being used in industries outside of healthcare. In this approach, data are explored without a specific hypothesis being established, relying only on a general sense that the data might reveal insights. Data mining is a subfield of computer science that uses algorithms to discover patterns of data interactions in large data sets. It uses artificial intelligence machine learning, classical statistics, and advanced database systems such as Hadoop. Examples of data mining tools are clustering and text mining. Cognitive computing tools such as IBM's Watson also support data mining.

Clustering

Clustering places objects into groups, or clusters, suggested by the nature of the data. The objects in each cluster tend to be similar to each other in some sense, and objects in different clusters tend to be dissimilar. If obvious clusters or groupings are developed prior to the analysis, the clustering analysis can be performed by simply sorting the data.

The clustering methods perform disjoint cluster analysis on the basis of Euclidean distances computed from one or more quantitative variables and seeds that are generated and updated by the algorithm. The user can specify the clustering criterion used to measure the distance between data observations and seeds. The observations are divided into clusters so that every observation belongs to at most one cluster.

EXHIBIT 8.9
Cluster Analysis of Sample Medicare Data



Note: Data used in this exhibit are the same as those used for the decision tree in exhibit 8.3.

After clustering is performed, the characteristics of the clusters can be examined graphically using a clustering package in software such R or SAS statistical packages. Exhibit 8.9 is a cluster analysis of the same Medicare data used for the decision tree in exhibit 8.3. Note that beneficiaries with chronic conditions cluster together because of their high use of inpatient services.

Text Mining

EHRs contain a significant amount of text, such as doctors' and nurses' notes. Therefore, a useful subset of data mining tools for healthcare providers is text miners. The case study that follows demonstrates the applicability of text mining to public health initiatives.

Case Example: Text Mining at the State Fair

The authors undertook an engagement in 2015 to assist a local nonprofit, Health Fair 11, an annual event sponsored by a local television station in Minneapolis–St. Paul in conjunction with the Minnesota State Fair (for more information, visit https://www.kare11.com/article/news/health/healthfair-11/a-healthy-minnesota-state-fair-tradition/89-296307902). The initiative

provides fairgoers with access to medical workers who check pulses, blood pressures, glucose levels, weight, and eyes and ears for potential health problems. Flu shots are also available, and advocacy groups are on hand to share health information on topics from gluten-free diets to stroke prevention to memory loss. Vendor groups include nonprofit organizations, professional associations, and for-profit companies.

The managers of Health Fair 11 were interested to know if their operational strategy needed adjustment. They surveyed a sample of 351 participants over six days. One of the key questions we asked was, "Why did you choose to get health screening at the state fair?" The general hypothesis was that the reason fairgoers used the Health Fair 11 screening services was either low cost or convenience. In addition to the results from these two options on our data collection form, we collected text answers (comments written freehand on the form).

Next, the managers engaged researchers who used the tools in SAS's text miner Topic to cluster the text responses. Exhibit 8.10 is the clustered response. Much to the investigators' surprise, the word *fun* appeared frequently. This unexpected result allowed the managers to pursue this concept with the organization and its vendors. They came to understand that the fairgoers felt empowered and engaged in this screening, as they were in control and did not have to go through the many gatekeepers of the traditional health system. This finding has proved useful to Health Fair 11 and carries important implications for primary care and population health.

EXHIBIT 8.10 Text Clustering Results from Health Fair 11 Survey

	W3 - Why did you get screening here?	
	Topic	**No. of documents**
1	fun,+learn,fun-check,doctor,doc's office	5
2	+screening,clinic,+check,office,health assessment	6
3	md,md's office,fair,offer,sucha	1
4	+learn,+live,fun-check,doctor,doc's office	3
5	time,fun-check,doctor,doc's office,doc	2
6	fair,information,valuable-love,access,convenient	2
7	+check,work,industry,+thing,health	4
8	doctor,fun-check,+visit,test,doc's office	2
9	sitting,down,cool,fan,fun-check	1
10	health assessment,keep,assessment,awareness,+build	2
11	random check,random,check,fun-check,doctor	1
12	doc's office,doc,office,+screening,work	3

Cognitive Computing for Data Mining

As discussed earlier, a major challenge for the analyst is data preparation and deployment of the analytical tools in the most sophisticated software packages. To address this issue, a number of technology firms are developing cognitive computing systems to simplify this work. Cognitive computing systems are designed to mimic human thought and provide natural language interfaces. A leading example is IBM Watson Analytics. Users load data into the system, and Watson performs significant preprocessing to suggest interesting correlations for the analyst to examine.

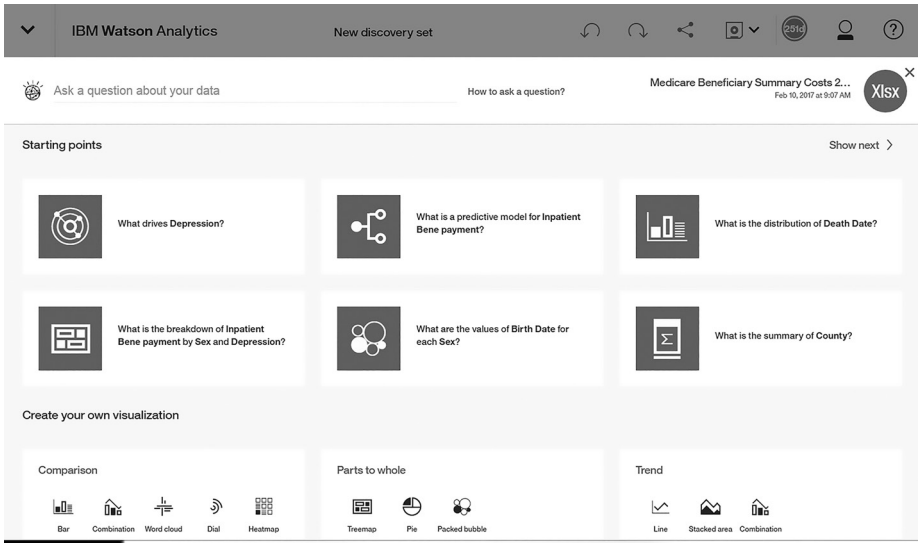
Exhibit 8.11 shows the starting screen from Watson as it looks at the Medicare beneficiary data used in earlier examples. It immediately offers six questions for the analyst to pursue. It also provides a natural language inquiry interface to delve deeper into the data.

Watson is a sophisticated example of a supervised learning tool and will continue to evolve as its underlying artificial intelligence software improves.

Conclusion

Analytics has become increasingly prevalent in healthcare. Hospitals and healthcare systems are using analytics as a means to gain insights into strategic, operational, and clinical issues. Today, the technology enables healthcare analytics to produce better visuals, build more sophisticated models, and analyze much more complex large data sets than at any time in the past. When

EXHIBIT 8.11 Opening Page Screenshot from Watson Analytics



Source: IBM Watson Analytics. Used with permission.

executed correctly, data are converted into actionable insights that allow enhanced decision making.

Discussion Questions

1. Identify a healthcare operating issue that could benefit from each of the analytical techniques:
   a. Descriptive
   b. Predictive
   c. Prescriptive
2. How could text mining be used to improve the care of patients with chronic disease?
3. Design a dashboard for each of the following care delivery types:
   a. Inpatient intensive care unit
   b. Outpatient imaging center
   c. Dental office
   d. Home health agency

Note

1. Portions of this section are adapted from BrightPoint Consulting (Gonzalez 2019). Used with permission.

References

Adams, D. A., R. A. Jajosky, U. Ajani, J. Kriseman, P. Sharp, D. H. Onweh, A. W. Schley, W. J. Anderson, A. Grigoryan, A. E. Aranas, M. S. Wodajo, and J. P. Abellera. 2014. "Summary of Notifiable Diseases—United States, 2012." *Morbidity and Mortality Weekly Report*. Published September 19. www.cdc.gov/mmwr/preview/mmwrhtml/mm6153a1.htm.

Barton, M. 2021. "Understanding Population Health Management: A Diabetes Example." HealthCatalyst. Published June 1. www.healthcatalyst.com/managing-diabetes-population-health-management.

Blumenthal, D., E. Malphrus, and J. M. McGinnis (eds.), Committee on Core Metrics for Better Health at Lower Cost, Institute of Medicine. 2015. *Vital Signs: Core Metrics for Health and Health Care Progress*. Washington, DC: National Academies Press.

Centers for Disease Control and Prevention (CDC). 2021. "COVID Data Tracker." Accessed July 12. https://covid.cdc.gov/covid-data-tracker/#datatracker-home.

———. 2009. "The Power of Prevention: Chronic Disease . . . the Public Health Challenge of the 21st Century." Accessed August 31, 2016. www.cdc.gov/chronicdisease/pdf/2009-power-of-prevention.pdf.

Cook, L. 2015. "America's Problem with Diabetes, in One Map." *US News & World Report*. Published April 9. www.usnews.com/news/blogs/data-mine/2015/04/09/americas-problem-with-diabetes-in-one-map.

Corbett, T. 2014. "An Innovative Approach to Improving Healthcare Outcomes." Published July 21. *Tableau*. www.tableau.com/about/blog/2014/7/Improving-Healthcare-Outcomes.

Davenport, T. H., and J. G. Harris. 2007. *Competing on Analytics: The New Science of Winning*. Boston: Harvard Business Review Press.

Gonzalez, T. 2019. "Dashboard Design: Key Performance Indicators and Metrics." BrightPoint Consulting. Published July 22. https://towardsdatascience.com/dashboard-design-key-performance-indicators-and-metrics-2b13745f5b2f.

Health Catalyst Editors. 2021. "Data Science Reveals Patients at Risk for Adverse Outcomes due to COVID-19 Care Disruptions." Health Catalyst. Published February 16. www.healthcatalyst.com/insights/healthcare-data-science-reveals-high-risk-patients.

Juran, J. M., and J. A. De Feo. 2010. *Juran's Quality Handbook: The Complete Guide to Performance Excellence*, 6th ed. New York: McGraw-Hill Education.

Kindig, D., and G. Stoddart. 2003. "What Is Population Health?" *American Journal of Public Health* 93 (3): 380–83.

Oxford Living Dictionaries. 2016. "Analytics." Accessed December 30. https://en.oxforddictionaries.com/definition/analytics.

Winters-Miner, L. A. 2014. "Seven Ways Predictive Analytics Can Improve Healthcare." Elsevier Connect. Published October 6. www.elsevier.com/connect/seven-ways-predictive-analytics-can-improve-healthcare.