EPRI | ELECTRIC POWER RESEARCH INSTITUTE

Innovative Applications of Customer Data Analytics: A Case Study

2015 TECHNICAL UPDATE

Innovative Applications of Customer Data Analytics: A Case Study

EPRI Project Managers
E. Petrill
C. Holmes

ELECTRIC POWER RESEARCH INSTITUTE

3420 Hillview Avenue
Palo Alto, CA 94304-1338
USA

PO Box 10412
Palo Alto, CA 94303-0813
USA

800.313.3774
650.855.2121

askepri@epri.com

www.epri.com

3002005720
Final Report, December 2015

DISCLAIMER OF WARRANTIES AND LIMITATION OF LIABILITIES

THE FOLLOWING ORGANIZATION(S), UNDER CONTRACT TO EPRI, PREPARED THIS REPORT:

Donnelly Communications

NOTE

For further information about EPRI, call the EPRI Customer Assistance Center at 800.313.3774 or e-mail askepri@epri.com.

Acknowledgments

The following organizations prepared this report:

Donnelly Communications
96 Amesport Landing
Half Moon Bay, CA 94019

Principal Investigator
E. Donnelly

Electric Power Research Institute (EPRI)
3420 Hillview Avenue
Palo Alto, CA 94304

Principal Investigators
E. Petrill
C. Holmes

This report describes research sponsored by EPRI. All of the figures in this report are from a September 2015 webcast entitled "Big Data in the Utility Industry." The webcast was hosted by Teradata and the presenters were Timotej Gavrilovic and Colin Kerrigan from PG&E.

EPRI would like to thank the following individuals at PG&E for providing information about the company's data analytics activities: Jan Berman, Grace He, Kevin Johnson, Chad Daniels, and Jen Owens Garza. In addition, the efforts of several analysts and other PG&E staff contributed to the work described in this report: Colin Kerrigan, Tingwen Bao, Pawel Zawadzki, Anne-Lise Laurain, Erik Gensler, and Timotej Gavrilovic.

EPRI was pleased to work with PG&E on this case study and appreciates how eager PG&E was to document its data analytics journey so that other utilities might benefit from PG&E's challenges and lessons learned.

This publication is a corporate document that should be cited in the literature in the following manner:

Innovative Applications of Customer Data Analytics: A Case Study. EPRI, Palo Alto, CA: 2015.

3002005720.

Abstract

Many utilities have invested in advanced metering infrastructure (AMI) over the last several years. AMI is an integrated system of smart meters, communications networks, and data management systems that enables two-way communication between utilities and customers. AMI provides opportunities for utilities to improve their business operations and expand the range of services they offer to their customers.

As utilities collect exponentially greater amounts of data on how consumers use electricity, they also face the challenge of how to use these data to gain a better understanding of their customers. This report is a case study of Pacific Gas and Electric Company's (PG&E's) customer data analytics journey. PG&E had deployed more than 5.1 million AMI devices as of September 2014—more than any other utility in the United States. PG&E also made the decision to develop a strong data analytics capability in-house. This report addresses the drivers for creating a centralized customer-side data analytics team, the forming and implementation of the team, and PG&E's data analytics platform and the systems used to support it. In addition, the report discusses examples of two current applications and provides a snapshot of three future data analytics projects. Finally, the report offers insights into some of the challenges PG&E faced—and continues to face—and the lessons learned as the company has expanded the scope of its data analytics activities.

Keywords

Advanced metering infrastructure (AMI)
Customer data analytics
Demand-side management (DSM)
Energy efficiency

EPRI | ELECTRIC POWER RESEARCH INSTITUTE

EXECUTIVE SUMMARY

Product ID: 3002005720

Innovative Applications of Customer Data Analytics: A Case Study

---

PRIMARY AUDIENCE: Electric utilities that want to understand how customer data analytics can improve their ability to understand customer energy usage and needs.
SECONDARY AUDIENCE: Organizations within electric utilities that have little experience with customer data analytics and are interested in expanding their analytics efforts beyond utility operations.

KEY RESEARCH QUESTION

As utilities collect exponentially greater amounts of data on how consumers use electricity, they also face the challenge of how to use this data to gain a better understanding of their customers. Customer data analytics can help utilities make use of this data. How does an electric utility create a centralized data analytics capability in their organization? What are the steps involved and what challenges might utilities face? What are some of the applications and what are the benefits?

RESEARCH OVERVIEW

This report is a case study of Pacific Gas and Electric Company's (PG&E's) customer data analytics journey. PG&E had deployed more than 5.1 million AMI devices as of September 2014, more than any other utility in the United States. PG&E also made the decision to develop a strong data analytics capability in-house. This report addresses the drivers for creating a centralized customer-side data analytics team, the forming and implementation of the team and PG&E's data analytics platform and the systems used to support it. In addition, the report discusses examples of two current applications and a snapshot of three future data analytics projects. Finally, the report provides insights into some of the challenges PG&E faced and continues to face and the lessons learned as the company has expanded the scope of its data analytics activities.

KEY FINDINGS

- Using load shapes constructed from AMI interval data combined with additional customer data enables better targeting of customers for specific utility programs, such as PG&E's SmartAC™ Program.
- Analyzing aggregated asset-level load distribution curves to determine where energy efficiency and demand response programs would have the most impact can help optimize distribution planning.
- An executive sponsor or company leadership that understands the value of data analytics is a key success factor.
- Once a data analytics team has been formed, build internal credibility with a few well-defined projects with clear goals and outcomes.
- The data analytics team should help define analytics projects and should be able to focus on projects that enable decision making.

VALUE STATEMENT

This case study describes how a large, complex electric utility created a centralized data analytics capability in its organization. It explains the steps involved, the challenges faced and lessons learned, and provides examples of data analytics projects that have resulted in value to the organization. Other utilities can use this information to facilitate discussions about the value that customer data analytics might provide to their own organizations.

This publication is a corporate document that should be cited in the literature in the following manner:

Innovative Applications of Customer Data Analytics: A Case Study. EPRI, Palo Alto, CA: 2015. 3002005720.

ELECTRIC POWER RESEARCH INSTITUTE

EXECUTIVE SUMMARY

HOW TO APPLY RESULTS

Utilities that are exploring the role of data analytics in their organizations can use the information in this report to develop a plan for creating and implementing their own data analytics team(s). They can also review the challenges and lessons learned presented in Section 5 to ensure the process for implementing a data analytics team is effective and brings value to the organization.

EPRI CONTACTS: Ellen Petrill, Senior Program Manager, 650-855-8939, epetrill@epri.com
Chris Holmes, Principal Technical Leader, 865-218-8116, cholmes@epri.com

Electric Power Research Institute
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com

Table of Contents

List of Figures

List of Tables

Section 1: Introduction

Many utilities have invested in advanced metering infrastructure (AMI) over the last several years. AMI is an integrated system of smart meters, communications networks, and data management systems that enables two-way communication between utilities and customers. A report from the Edison Foundation Institute for Electric Innovation found that as of July 2014, over 50 million smart meters had been deployed in the United States. Close to 90% of these meters are residential, covering over 43% of U.S. homes. [1].

Advanced metering infrastructure (AMI) has been adopted by many utilities and provides opportunities to improve business operations and offer additional services to electricity customers.

AMI provides opportunities for utilities to improve their business operations and expand the range of services they offer to their customers. For example, utilities are integrating AMI systems with outage management systems and distribution management systems to help manage outages and improve distribution system monitoring. AMI also enables services such as automated budget assistance and bill management tools, energy use notifications and smart pricing and demand response programs.

As utilities collect exponentially greater amounts of data on how consumers use electricity, they also face the challenge of how to use this data to gain a better understanding of their customers. EPRI is striving to help its member utilities address this challenge. Many vendors offer products from standalone custom analysis tools to enterprise-wide integrated platforms, but utilities need guidance about which products will work best for their business model. A 2014 EPRI report compiled and reviewed sources of customer data analytics products [2]. Another option for utilities is to analyze the data themselves, but collecting, managing, and analyzing this data requires advanced analytics capabilities that few utilities currently possess.

This report is a case study of Pacific Gas and Electric Company's (PG&E's) customer data analytics journey. PG&E had deployed more than 5.1 million AMI devices as of September 2014, more than any other utility in the United States [3]. PG&E also made the decision to develop a strong data analytics capability in-house. This report addresses the drivers for creating a centralized customer-side data analytics team, the forming and implementation of the team and PG&E's data analytics platform and the systems used to support it. In addition, the report discusses examples of two current applications: targeting potential customers for PG&E's SmartAC™ program, and a Targeted Demand Side Management (DSM) program to better serve customers and to help prioritize distribution upgrades. The report also provides a snapshot of three future data analytics projects: enhancement of the Targeted DSM program, segmentation of PG&E's small and medium business (SMB) customers and a

project to develop a 360-degree view of all of its customers by collecting and analyzing dozens of individual customer attributes. Finally, the report provides insights into some of the challenges PG&E faced and continues to face and the lessons learned as the company has expanded the scope of its data analytics activities.

Section 2: PG&E's Customer-Side Data Analytics Team

This section provides an overview of PG&E, a brief history of the drivers and implementation of the company's customer-side data analytics team, and a description of PG&E's data analytics platform.

PG&E Overview

PG&E provides electric services to over 5 million electric customer accounts and has a 70,000 square mile service territory.

PG&E is one of the largest combination gas and electric utilities in the United States. The company provides electric and natural gas services to approximately 16 million people and has approximately 5.4 million electric customer accounts and 4.3 million natural gas customer accounts. PG&E's 70,000 square mile service territory stretches from Eureka in the north to Bakersfield in the south, and from the Pacific Ocean in the west to the Sierra Nevada in the east [4], as shown in Figure 2-1. PG&E is headquartered in San Francisco and is the electricity service provider to many world-class technology companies, including Oracle, Apple, Google and Facebook, to name just a few. The company has approximately 23,000 employees.

Figure 2-1
PG&E's service territory

PG&E is one of three regulated investor-owned electric utilities in California—along with Southern California Edison and San Diego Gas & Electric—and is regulated by the California Public Utilities Commission (CPUC). California is also one of 29 states to have a Renewable Portfolio Standard (RPS). The state's RPS mandates that by 2020, 33% of electricity generated must come from renewable energy sources [5]. California Senate Bill 350, signed in October 2015, extends the timeline to 2030 and the requirement to 50% of retail sales that must come from renewable energy. Soon these changes will be implemented [6]. In addition, California's regulatory structure decouples profits from energy sales revenue and provides incentives for demand reductions and energy efficiency.

PG&E has deployed over 5 million AMI devices as part of its SmartMeter™ program.

In the last several years PG&E has embarked on multiple initiatives to work with customers to improve energy efficiency (EE), provide incentives and technical assistance to customers investing in energy management technologies that also enable them to perform demand response (DR), and shift its generation portfolio to incorporate more renewable energy. As mentioned in the previous section, as of September 2014 PG&E had deployed more than 5.1 million AMI devices as part of its SmartMeter™ program. PG&E's SmartMeter™ system collects electric and natural gas usage data from its customers' homes and businesses. SmartMeter™ electric meters record residential electric usage hourly and commercial electric usage in 15-minute increments [7].

Formation of Data Analytics Team

PG&E wanted to improve the tracking of its energy efficiency and demand response programs.

PG&E recognized that with the data collected by the SmartMeter™ system, it could improve how it tracks the performance of its EE and DR programs. The company could also improve how it conveyed this information to PG&E's leadership. Executives would benefit from data in a format that was concise and visual so they could easily grasp the essentials and use the information to improve business operations. Although the company had some data analytics capabilities, this expertise was scattered throughout PG&E's Customer Energy Solution (CES) organization. An analysis requiring data analytics typically took up to two to three weeks, and the analysts had to collect data manually and transfer large amounts of data between legacy systems across organizational boundaries. Reporting was typically done in Excel. PG&E's historic approach to data analytics is shown in Figure 2-2.

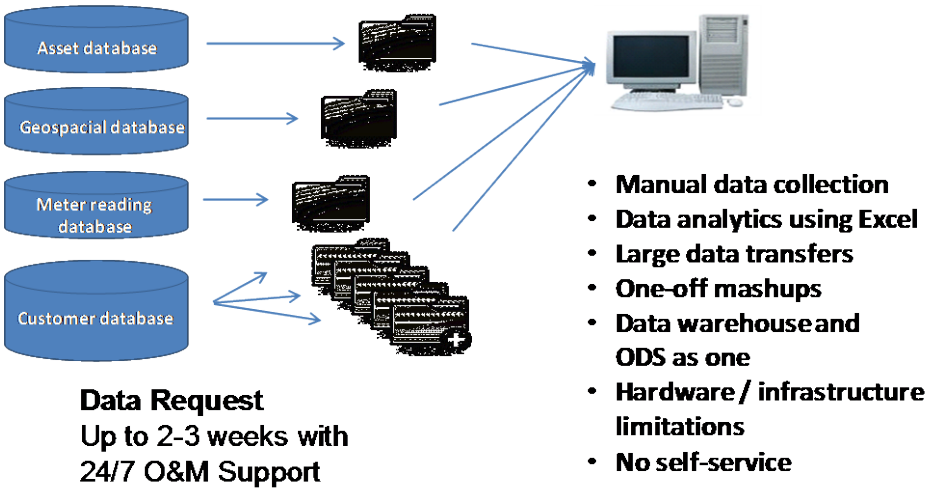


Figure 2-2
PG&E's Historic Approach to Data Analytics

In 2013, PG&E combined two data analytics teams and a group responsible for regulatory reporting to form a customer-side data analytics team.

Approximately four years ago, two data analytics teams were formed in the CES organization, one of three organizations in the Customer Care group. Originally the two teams had different responsibilities. One team focused on performing simple analyses and creating visualization dashboards. The second team focused on complex analytics projects, and some of the analyses they performed were passed on to the other team to create visualization projects. In 2013, the two teams were combined along with a group responsible for regulatory reporting to form one customer-side data analytics team. The team currently has 27 members and is a shared data and analytics resource for PG&E's Customer Care organization, where it functionally resides, along with other internal clients. The chart in Figure 2-3 shows where the customer-side data analytics team sits in the Customer Care organization. The chart does not include all of the other organizations under the Customer Service and Customer Operations areas; the blue boxes represent the reporting structure for this team only.

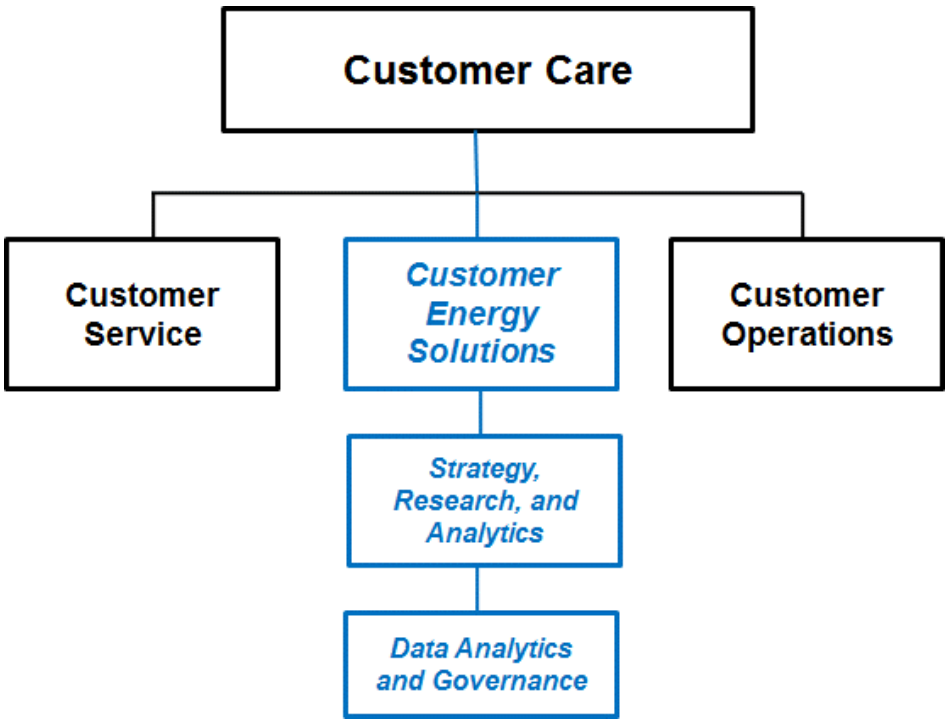


Figure 2-3
Organization Chart Showing Customer-Side Data Analytics Team

The customer-side data analytics team provides strategic and analytical support to many different organizations within PG&E.

The team's mission is to provide strategic and analytical support for decision making regarding energy efficiency, demand response, distributed generation, pricing products and electric vehicles. They work consistently with several different organizations through the company—including Electric Operations and Energy Procurement—and provide support as needed to many other organizations. In November 2015 the team was reorganized into three functional areas to improve its efficiency: a functional engagement area that partners with internal clients as ad hoc consultants; a solutions development area that creates analytical and visualization solutions to address specific business needs; and a data reports and platforms area that conducts data queries and reporting and oversees platform development. Figure 2-4 provides an overview of the data analytics team responsibilities.

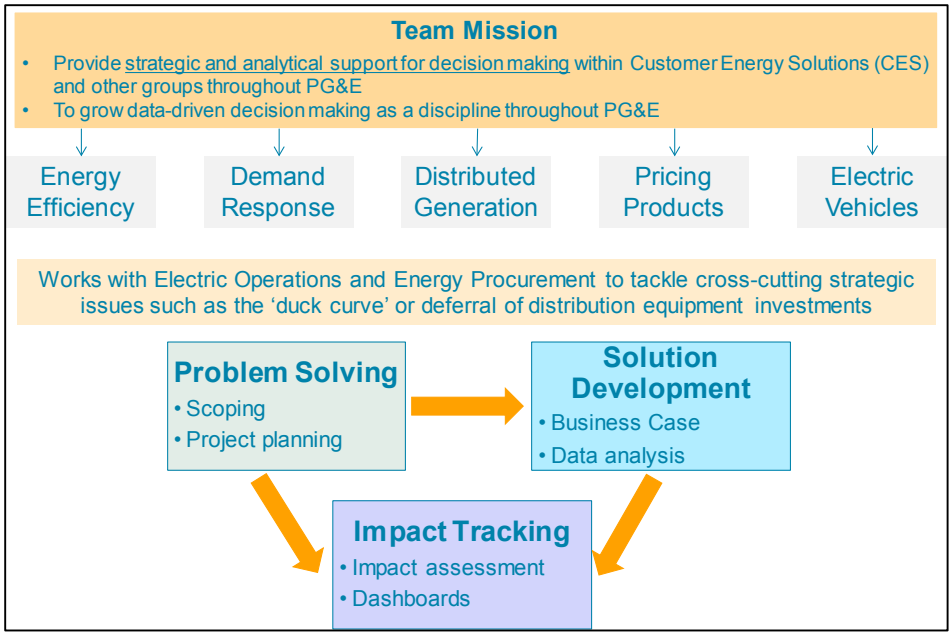


Figure 2-4
Overview of Data Analytics Team

As the team has expanded over the last four years, the role of analytics at PG&E has also expanded as the company recognized that integrated data analysis could enable decision making, penetrate organizational silos and drive the company's long-term strategy. The evolution of PG&E's analytics capabilities is shown in Figure 2-5.

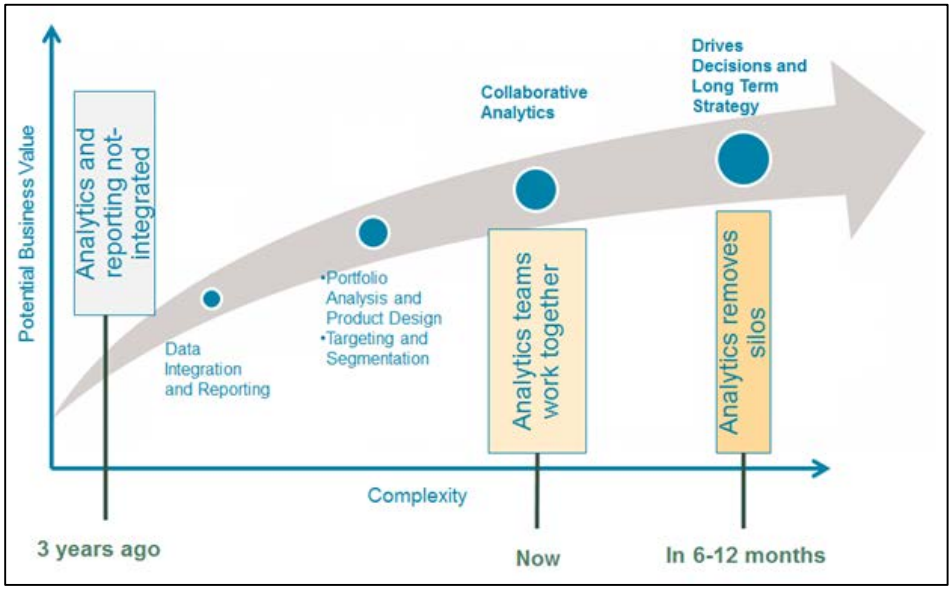


Figure 2-5
Analytics Development at PG&E

PG&E has several analytics teams throughout the company. A Data Users Committee of Practice has been formed with hundreds of members and several dozen meet quarterly to connect and learn about each other's work.

An ideal data analytics team would have a skill set that includes statistics, mathematics, machine learning, cloud computing, consulting, strategy development, policy and market understanding, data management, visualization techniques and engineering. The team would also have expertise in multiple analytical tools, programming languages, and platforms; SAS, SQL, Tableau, R and Python are some of the central tools in PG&E's data ecosystem.

Integrated Data Analytics Environment

PG&E invested in an integrated data analytics environment to take advantage of interval data being provided by its AMI.

PG&E analytics capabilities depend on both its staff expertise and its integrated data analytics (IDA) environment. The original impetus for investing in an IDA environment was that multiple internal teams wanted to take advantage of the information being provided by interval data for a variety of use cases. In addition, the CPUC, the California Energy Commission (CEC), and other external parties were submitting continuously increasing requests for customers' energy usage information. Manual extraction and integration of information from PG&E's multiple databases would have required hiring additional staff and would have also created non-compliance and error risk from the labor-intensive processes involved. A net present value analysis showed that it was more cost-effective for PG&E to invest in an IDA environment rather than expand processes which had not been designed for scale.

At the heart of the IDA environment is a database from Teradata, a leader in data analytics solutions and services. Teradata houses the interval data for all AMI accounts in PG&E's service territory. Teradata also has more than 100 technology partners, including PG&E [8].

SQL tools provide a mechanism for querying the Teradata system as well as PG&E's internal customer data and other demographic data, which are consolidated into an internal Customer Data Warehouse (CDW).The CDW serves to compile data from systems such as its SmartMeter™ system, the Customer Care and Billing (CC&B) system containing customer data, the Advanced Billing System (ABS) and energy efficiency audits and databases that track energy efficiency and demand response program participation, and make the data available for reporting and analytics.

Demographic data from external partners is integrated into the CDW. PG&E also uses statistical analysis software from SAS to help collect and analyze information for different internal client use cases. Finally, PG&E uses Tableau's visualization software to create "dashboards" that convey specific information in a format that is easy to understand and helps enable decision making.

Section 3: PG&E Customer-Side Analytics Project Examples

This section provides two examples of current PG&E analytics projects that are providing value to the company: targeting potential customers for the SmartAC™ program, and a Targeted DSM program to help prioritize distribution upgrades.

Targeting Potential Customers for the SmartAC™ Program

The goal of PG&E's SmartAC™ program is to smooth out energy usage peaks to ensure a reliable energy supply on days when peak demand is high. This typically occurs on hot days due to high use of air conditioning. The program entails PG&E installing a device on or near a participating residential customer's air conditioner. If a local or state energy emergency is anticipated, the device will be remotely activated and the customer's central air conditioner compressor will cycle on and off in short increments totaling about 15 minutes of every half hour. Most customers will not see a reduction in their energy bills, but PG&E offers a $50 incentive to customers for enrolling in the program [9].

Challenge

PG&E wanted to improve how it targeted customers for the company's SmartAC™ program.

Targeting potential customers for the SmartAC™ program requires a sophisticated marketing approach to find customers with air conditioning load available to cycle. PG&E had historically used a propensity model to identify likely participants, but with only monthly energy usage, PG&E had no way of knowing how much electricity customers were using on an hourly or even daily basis. Thus, it was difficult to determine whether customers had air conditioners or were using them during coincident peak hours. Due to these factors, PG&E's initial targeting efforts for the SmartAC™ program were not as effective as they could be.

Solution

The analytics team believed that improving targeting for the SmartAC™ program could allow the program to be more effective. PG&E's SmartMeter™ system provides interval data on an hourly basis that allows PG&E to track customer's energy use throughout the day and create load shape profiles of potential participants in the SmartAC™ program. Analysts can examine

customer energy use profiles and electricity load shapes to identify likely customers for the SmartAC™ program.

Load shapes helped PG&E gain a better understanding of customers' hourly energy usage.

Figure 3-1 provides a "before and after" look at how the analysts improved PG&E's targeting for the SmartAC™ program. As shown in the figure, two customers were considered to be strong candidates for the SmartAC™ program based on their daily usage, as shown on the left. The data for the table on the left was obtained from marketing propensity scoring provided by a third party, demographic data, and daily energy usage data. Analysts used interval data provided by PG&E's SmartMeter™ system to create a load shape for each customer and were therefore able to show that Customer 1's peak energy usage occurred in the evening, while Customer 2's was in the morning, making Customer 1 the better target for the SmartAC program™ since air conditioning is more likely to be used in the evening than in the morning on hot days.

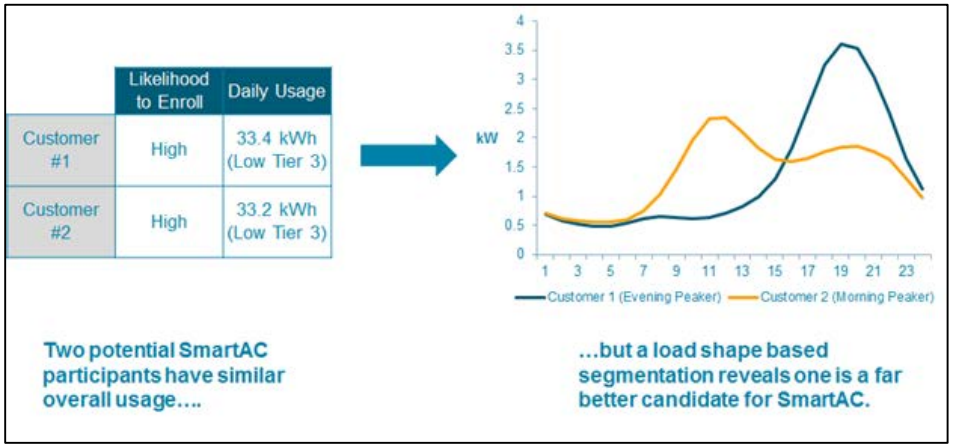


Figure 3-1
Using load data to understand customer energy use

The analytics team applied the load segmentation to all residential customers, which showed how extreme the concentrated load resource was among a subset of customers. As shown in Figure 3-2, 40% of residential customers account for over 80% of the load resource potential for the SmartAC™ program.

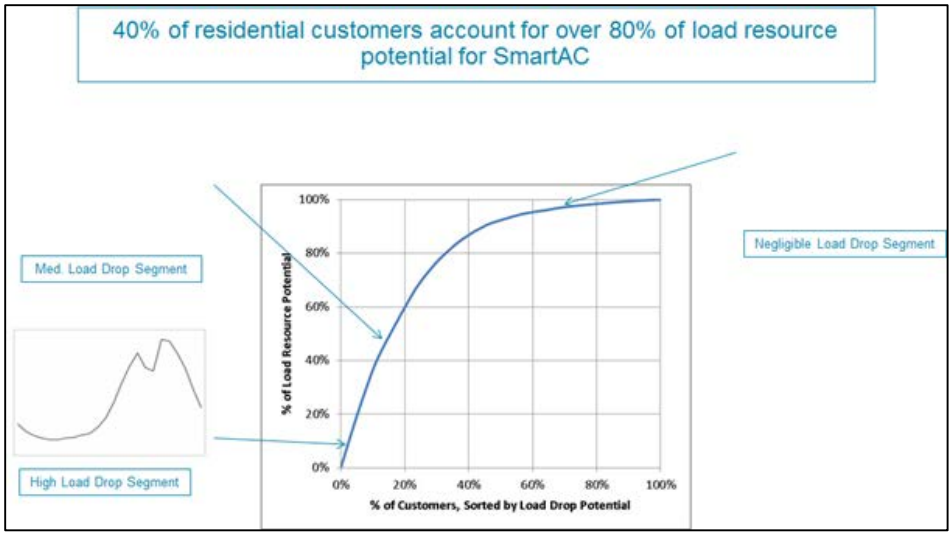


Figure 3-2
Load segmentation shows which customers are the best targets for SmartAC™

The analysis of customer load segmentation data to improve targeting for PG&E's SmartAC™ program was one of the first successes of the customer-side data analytics team.

When the analysis was completed, the analytics team was able to provide PG&E sales staff with a targeted list of customers who would likely participate in the SmartAC™ program, and whose electrical load provided a good opportunity for load reduction from SmartAC™ participation. This project was one of the first successes of the analytics team and provided a compelling case to PG&E management for the value of integrating multiple data points, in this case to develop targeted marketing for a PG&E program. In addition, the load shape analysis included algorithms to analyze why some SmartAC™ devices were no longer working. As a result, PG&E reduced service technician cost, because the technicians didn't have to travel as often to customers' homes to examine the devices to determine why they were malfunctioning.

Targeted Demand Side Management (DSM) for Electricity Reliability

With adoption of distributed energy resources (DER)—for example rooftop solar energy incented through the California Solar Initiative—California regulators and utilities face the challenge of ensuring that the reliability of the distribution grid will not be adversely affected. DERs include distributed renewable generation, energy storage, electric vehicles, energy efficiency and demand response. If approached intelligently, DERs can potentially be deployed to address system needs identified in the planning process. The availability of granular operating and customer data presents an opportunity to test whether targeted deployment of DERs can provide value to the distribution grid.

Challenge

PG&E's desire to better integrate planning between its customer and grid operations has resulted in a pilot to test whether targeted deployment of demand side resources can provide grid value. To do this, the company needed to be able to deliver peak demand reductions and/or distributed generation in the right

PG&E wanted to deploy energy efficiency or demand response programs in a way that would help meet specific grid distribution needs.

quantity, in the right place and at the right time to meet grid operational needs [10].

Historically PG&E had not had the systems or an integrated data management strategy that enabled gathering sufficient data to deploy energy efficiency or demand response in a way to meet specific distribution grid needs. This made it difficult to calculate energy benefits (energy and demand response savings) vs. non-energy benefits (such as improved reliability and reduced carrying cost of capital). It was also difficult for PG&E to target likely customers because they lacked an understanding of which customers were most likely to adopt energy efficiency and demand response offerings. Targeting customers effectively required information about their load and load profile and their participation in energy efficiency and demand response programs to date.

Solution

PG&E launched a pilot project to identify whether or not customers could be targeted with demand side management programs to provide sufficient coincident demand reduction to meet distribution needs. Four substations under consideration for upgrade were chosen for inclusion in the pilot, as shown in Figure 3-3. Specific information about each location is shown in Table 3-1.

Figure 3-3
Targeted DSM pilot focused on four substations

Table 3-1
Breakdown of PG&E Targeted DSM Pilot Locations

Substation #	Location	Customers*	Load Reduction Goal
1	Yuba City, northern CA	14,000 RES 1,000 SMB 400 LCI	2,000 kilowatt peak load reduction by the end of 2015
2	Tracy, Central Valley, CA	7,000 RES 250 SMB 250 LCI	2,500 kilowatt peak load reduction by the end of 2015
3	Jackson, Central Valley, CA	3,500 RES 700 SMB 150 LCI	800 kilowatt peak load reduction by the end of 2015
4	Fresno, Central Valley, CA	23,000 RES 2,000 SMB 500 LCI	2,500 kilowatt peak load reduction by the end of 2015

* RES = residential, SMB = small/medium business, LCI = large commercial/industrial

This effort required the analytics team to integrate information from several databases. These included: interval data from the SmartMeter™ system to determine customer loads during substation peaks; SmartAC™ data; census block data such as income, language spoken and other demographic information; billing data; the sales pipeline for energy efficiency programs for larger customers; and Evaluation, Measurement and Verification (EM&V) data on participation in energy efficiency or demand response programs.

The team analyzed aggregated asset-level load duration curves to help determine whether energy efficiency or demand response programs would be the best way to reduce energy use.

Aggregated asset-level load duration curves based on AMI data were analyzed based on year-round data. Figure 3-4 shows fictionalized data that is representative of actual trends. The figure shows fictionalized feeder load information for Substation 2 and demonstrates that at the local level, different feeders in the same substation can have very different load profiles. The figure shows two feeders on the same bank with two different customer types and their respective load duration curves. Load on feeder 9913 is composed of primarily industrial customers, who tend to have a relatively flat load duration curve. For an asset that shows a pattern of consistently high energy usage (top curve, Figure 3-4), energy efficiency products could help reduce the overall energy use on the targeted asset and potentially enable rescheduling of distribution upgrades at that location.

The load on feeder 9914 is composed of mostly residential customers with air conditioning. The load duration curve for this feeder shows a short period of high load (bottom curve, Figure 3-4.) To address this short high load, demand response programs would be more practical than energy efficiency programs, since demand response can reduce peak loads of short duration. Because the feeder 9914 customers' peak demand are driven by residential air conditioning use, SmartAC™ and other demand response programs would be effective offerings to these customers. Participation in the demand response programs could result in reducing the load during the top few hours of the year on that

feeder, delaying the need for upgrade by a year or more, meaning that PG&E could reschedule the distribution system upgrades in the area.

The data integration effort is ongoing and full-scale implementation may take several years, but initial results from the pilot are encouraging.

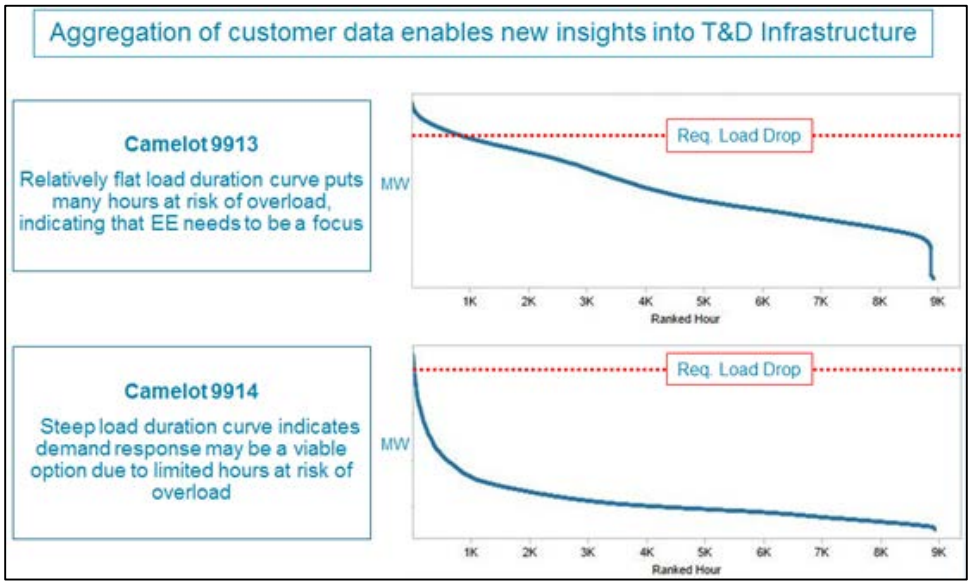


Figure 3-4
Fictionalized feeder load information for Substation 2

Another part of the pilot was to provide account representatives with information that would allow them to target specific customers for demand response programs that would have the most likelihood of reducing peak loads and therefore enable rescheduling of the distribution system upgrades. The analytics team created a Tableau dashboard with data extracted from the Teradata system and other systems, including SmartMeter™ data, energy efficiency and demand response program participation, demographic data and customer industry information. The dashboard shows how much load individual customers contribute to the distribution system during local peak-load times (peak-coincident demand), and their past participation in energy efficiency and demand response programs. Figure 3-5 shows a Targeted DSM dashboard containing fictionalized data representative of actual trends.

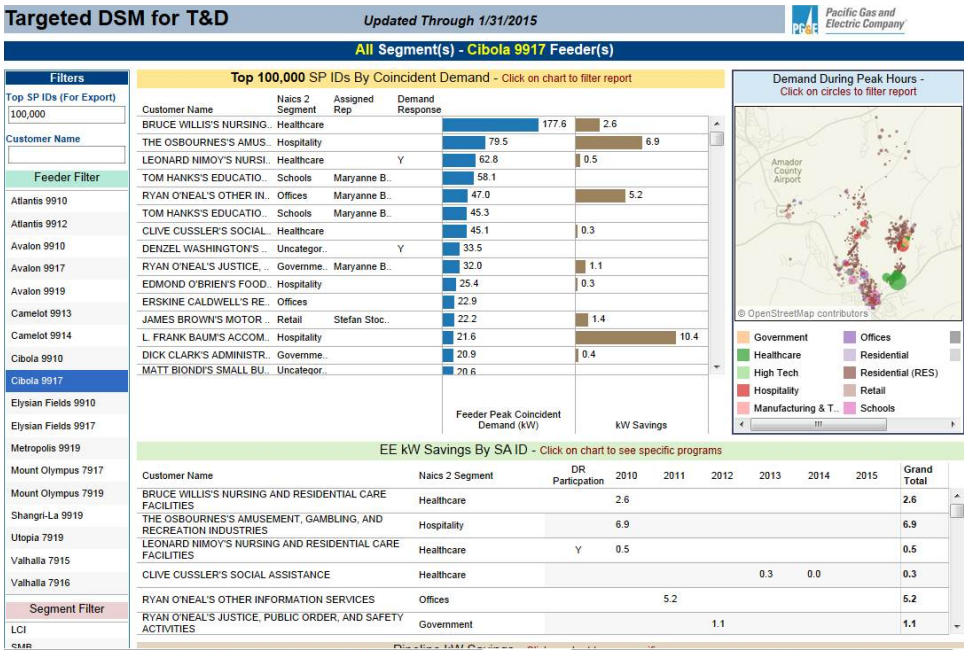


Figure 3-5
Targeted DSM Dashboard using fictionalized data representative of actual trends

The Targeted DSM pilot resulted in PG&E rescheduling planned capacity investments at the four substations and also demonstrated the potential benefits of a broader data integration strategy.

Before PG&E invested in the Teradata platform it was very difficult to tell how much individual customers were contributing to local infrastructure peaks and the analysis to determine where distribution upgrades could be reprioritized would have taken many months. PG&E plans to integrate the data shown in the dashboard into the distribution planning process and to continue using analytics and dashboards to integrate the information currently located in different databases, since dashboards provided an "at a glance" analysis that couldn't be otherwise achieved.

The Targeted DSM pilot and the data analytics conducted by the team provided significant value to PG&E and its customers. By allowing planned capacity investments at each of the substations to be rescheduled in the next three to five years, PG&E might save up to $2.5 million to spend on other high-priority projects [10]. In addition, the Targeted DSM pilot involved integrating enormous amounts of data across the organization, which helped highlight the potential benefits of a broader data integration strategy.

PG&E plans to roll out the pilot to an additional four locations in 2016 and to integrate the pilot results into the company's Electric Operations and Distribution Planning functions. They also plan to refine the TDSM dashboard, which is discussed in the next section of this report.

Section 4: PG&E Customer-Side Analytics Future Project Examples

This section provides three examples of planned future PG&E customer-side analytics projects: enhancement of the Targeted DSM program, segmentation of PG&E's small and medium business (SMB) customers and a project to develop a 360-degree view of all of its customers by collecting and analyzing dozens of individual customer attributes.

Enhancement of Targeted DSM Project and EPIC Project for Targeted Demand Reduction

PG&E will launch an additional pilot for targeted demand reduction through targeted data analytics to identify and test a scalable data integration and analytics platform that will help optimize locational deployment of energy efficiency, demand response and other programs and technologies.

As discussed earlier, PG&E plans to use the results of the Targeted DSM pilot to incorporate additional customer-level information into its distribution planning process. The next phase of the project is to apply the learning from 2015 to additional locations. The next step is to automate the data processing that supplies data for the Targeted DSM dashboard and augment the information provided in the dashboard.

Going forward, an additional pilot project was recently awarded funding through PG&E's Electric Program Investment Charge (EPIC) program. EPIC was created by the CPUC in 2011 to support investments in clean energy technologies that provide benefits to the electricity ratepayers of PG&E, San Diego Gas & Electric and Southern California Edison. The program is administered by the CEC and the three California electricity IOUs. The program funds clean energy research and demonstration and deployment projects that support California's energy policy goals and promote improved electricity reliability, lower costs and increased safety [11]. PG&E's recent EPIC research [12] funding includes a project that builds on the Targeted DSM work, and will test a new high-power data analytics platform to perform and extend the TDSM insights, make the analysis scalable to the full PG&E territory, and incorporate additional distributed energy resources, including distributed generation and electric vehicles, and include more data sources and cost-effectiveness information in the optimization.

The pilot project for Targeted Demand Reduction through Targeted Data Analytics is aimed at identifying and testing a scalable data integration and analytics platform that enables cross-platform analytics with very large datasets; the solution will involve optimizing locational deployment of EE, DR and other technologies and programs, building additional algorithms on the base-data and

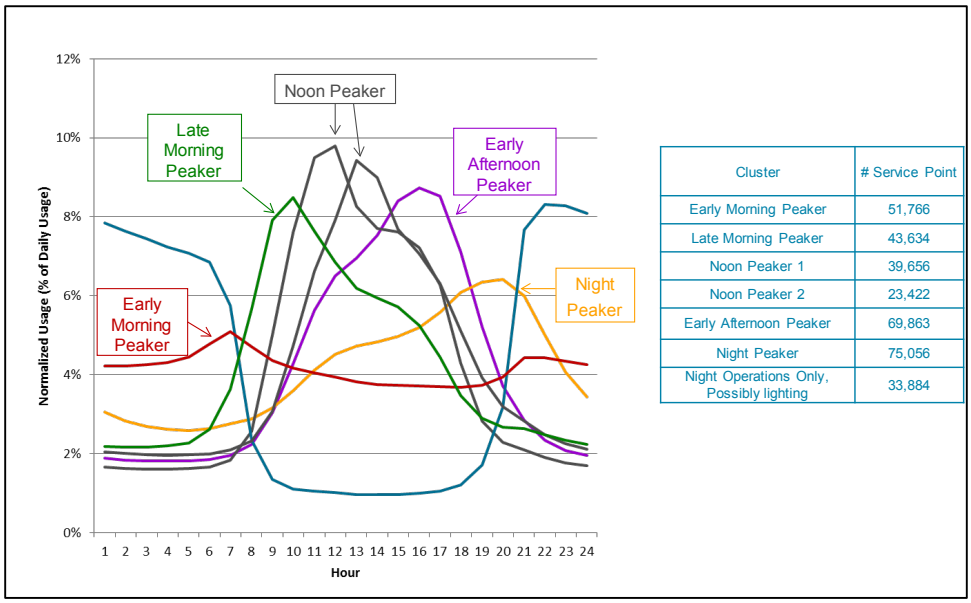
analytics used for the Targeted TDSM pilot. The team plans to apply advanced analytics and machine learning algorithms to screen the most cost-effective DSM deployments targeted to customers that can benefit, are likely to adopt and are in locations where the changes in energy usage will be most valuable to the grid. This analysis will be used to do a first pass screening in April 2016, once PG&E's Distribution Planning organization identifies the list of potential targeted assets, and to optimize and deploy the best mix of customer programs and technologies for the grid needs at that location. This will enhance the incorporation of DERs and customer load management in the distribution planning process, and help provide a more flexible, resilient and reliable electrical distribution system that will be able to incorporate the additional distributed generation that is anticipated in coming years.

Small and Medium Business (SMB) Segmentation

PG&E is using data analytics to improve segmentation of its small and medium business customers.

PG&E is in the process of improving its SMB segmentation by gaining a clearer picture of these customers' energy behavior. Servicing SMB customers is challenging for account representatives because there are a large number of customers spread throughout PG&E's service territory and it is difficult and expensive to reach them all. These customers also tend to be active seven days a week so PG&E has to keep this in mind when scheduling energy system upgrades, maintenance, or other services to avoid disrupting electricity service.

A preliminary analysis indicated that load shapes spread over 24 hours may be a useful segmentation approach compared to business type. As shown in Figure 4-1, there is a wide variety of load profiles among SMB customers, from "Early Morning Peakers" to "Night Peakers."



Cluster	# Service Point
Early Morning Peaker	51,766
Late Morning Peaker	43,634
Noon Peaker 1	39,656
Noon Peaker 2	23,422
Early Afternoon Peaker	69,863
Night Peaker	75,056
Night Operations Only, Possibly lighting	33,884

Figure 4-1
Preliminary SMB segmentation

The analytics team is developing a "DNA code" or "load fingerprint" of PG&E's SMB customers based on data provided by the company's SmartMeter™ system.

The analytics team is now working on developing a "DNA code" or load "fingerprint" for SMB customers based on customers' interval data characteristics, as provided by the SmartMeter™ system. The current approach for the DNA code identifies six characteristics:

1. Seasonality.
2. Weekly pattern.
3. Peak time.
4. Peak duration.
5. Volatility.
6. Base load.

As shown in the illustrative example in Figure 4-2, some SMB customers have a high summer load due to refrigeration and/or air conditioning, while others have distinctive load shapes revealing peak energy usage due to other end-uses. As the analytics team continues to refine this approach, PG&E account representatives will be able to make informed decisions about whether energy efficiency and conservation measures or demand response and load shifting would be the best approach for reducing SMB customers' energy usage. The analytics team is hoping to do the kind of advanced modeling about their customers being conducted by companies such as Amazon and Walmart, and believe that they may be the first utility to attempt this kind of analytics with such a large SmartMeter™ data source at the core.

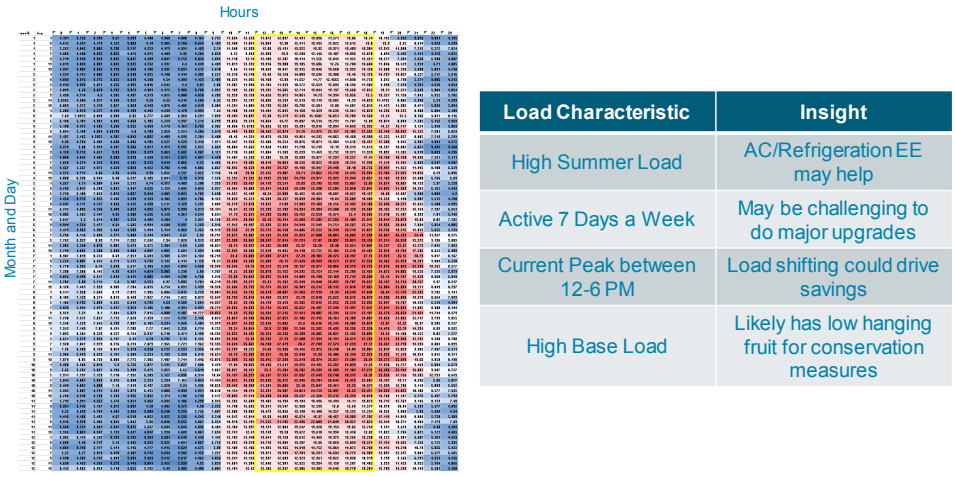


Figure 4-2
Illustrative example of SMB load segmentation

Customer 360

PG&E's Customer 360 project will identify and collect dozens of attributes for each customer.

The goal of this project is to further advance PG&E's understanding of all of its customers by identifying and collecting 40-50 attributes per customer to provide

a 360-degree view of each customer. PG&E's SmartMeters™ already provide significant amounts of information and the data analytics team plans to augment this data with information from its on-bill financing (OBF) system, from other demographic data sources. PG&E's OBF program can help increase adoption by financing the cost of implementing energy efficiency and demand response measures by allowing customers to spread the cost over time. Customers then pay for the cost of the equipment and the financing cost on their bills to PG&E.

Section 5: Challenges and Lessons Learned

The role of analytics and of the customer-side analytics team at PG&E has been evolving over the last several years. As PG&E continues to move toward integrating data analysis into its decision making and strategy development, the analytics team identified several challenges the company faces. The team also highlighted several lessons learned about implementing a data analytics capability into an organization that could be useful for other utilities.

Challenges

Key challenges for the customer-side data analytics team include a lack of a broad data integration strategy, not being involved in the early stages of projects, and hiring and retaining data analytics experts.

PG&E is a large, complex company and many of its organizational entities have their own databases and processes for collecting, analyzing and using data, including communicating results to the company's leadership. These organizational and data silos can make it difficult to provide the most effective results to company leadership, let alone for the analytics team to conduct the complex analyses they are often asked for.

In addition, when the team receives data for an analysis, it is often in a variety of formats and they have to spend considerable amounts of time addressing data format and quality before the data can be used for analysis. This lack of data integration results in some of the data being collected by PG&E from its various customer-facing systems not being utilized. The development of data input systems should take into consideration the end-uses of the data, such that the collection process, storage, and management of the data enable analysts access to as much of the collected data as possible. In this way, ensuring a usable data output system will allow more data to be used to inform and enable decision making.

Yet another challenge is that if performance data on various PG&E programs is not presented to the company's leadership in a useful and timely way, there is a risk of issues not being addressed. One team member said:

"If we can't report it swiftly and accurately, it never happened."

The team is working on developing a broader data integration strategy for the Customer Energy Solutions group to help address these issues.

Another challenge is that when the analytics team is asked to begin a new project, the team's internal customers do not always know exactly what data or

analyses are required to meet their objectives. In addition, since integrated data analyses are still fairly new at PG&E, the team's internal customers may not know what kinds of analyses are possible. If team members are not involved in the early discussions of the project, they cannot help define the proper scope and elements for an analysis. In addition, sometimes the team's internal clients tend to focus on what they individually see as most important and risk losing sight of the project's overall goal.

At the same time, the team struggles with defining its role. The utility industry is not nearly as advanced in terms of data analytics as many other industries. Since data analytics at PG&E had been historically focused on simple data reporting, the team wants to ensure they are spending the majority of their time on analytics projects that will provide insights to enable strategic decision making. These projects are the more complex ones that make the most of their expertise.

Finally, data analytics is a field that has grown exponentially over the last several years. There is fierce competition for data analytics experts, especially in the San Francisco Bay Area. As a result, PG&E has found it difficult to hire and retain data analysts.

Lessons Learned

The Role of an Executive Sponsor

An executive sponsor and a centralized data analytics function are key success factors.

It requires a significant undertaking to create and implement a centralized data analytics capability in any organization, especially in an organization that is part of a large, complex company like PG&E. In addition, the utility industry has lagged behind other industries in recognizing the value that data analytics can provide to the organization. It is very important to have an executive sponsor who understands the key role of data and analytics. If a company's leadership is not metric-driven, it is unlikely to recognize the importance of not only collecting data but integrating the data to enable rapid analytics that provide insights to facilitate decision making. An executive sponsor is key, as well as leaders who recognizes the value of metrics.

Centralizing the Data Analytics Function

Integrating and/or centralizing data analytics into one team or functional unit within the company—such as PG&E's Customer Energy Solutions group—is key to success. Proactively managing and using data to enable strategic decision making can help break down organizational silos and improve internal communication. Through its partnership with Teradata and in its work other industry leaders, PG&E has observed a variety of different ways that an analytics team can be organized. Regardless of a company's organizational structure, an analytics team needs to serve a broad internal client base within the company.

Making Data Accessible and Integrated for Business Users

One key development was the creation of IDA, allowing data analysts to tap interval data for different business users, and different purposes as use cases arose.

This solution also integrated other key data sources, and provided a foundation for expansion, while linking with the SAS analytics platform in order to provide value from the data.

Build Internal Credibility

Once a centralized analytics team is created, it is important to build internal credibility and for the team to be involved in the planning of new analytical projects.

As demonstrated by the analytics team's success with providing a targeted list of customers for the SmartAC™ program, identifying an application of data analytics that can provide early and clear value to the organization helps build credibility for the function and the team. While initial success is critical, so is continuing to expand the capabilities and outputs of the team, as demonstrated by the early success of the Targeted DSM project.

Help Define Analytics Projects

Once the internal credibility of the team has been established, the analytics team can continue to demonstrate its value to the organization by being involved in the initial discussions about new analytical projects. Their expertise can help shape the project and ensure that it has an actionable outcome rather than just being an exercise in data collection and/or analysis. In addition, including the analytics team when a project is being scoped will help improve internal communication and break down data and organizational silos.

Focus on High-Value Analytics Projects

Data analysts can provide the most value by focusing on projects that present results in a visual context that will enable decision making.

Simple reporting does not drive analytics or deliver much value. Analysts need to focus on projects that present the results in a visual context that can be easily understood and that facilitates decision making. These projects require a team that can visualize information and conduct complex analyses. These projects will provide the most value and will enable the company to operate more efficiently and better serve its customers.

Hire and Retain Analysts with Diverse Analytical Capabilities

The use of data analytics in many companies—including PG&E—is becoming common in many industries. In the San Francisco Bay Area where PG&E is headquartered, there are hundreds of companies employing data analysts. An organization that is focused on using customer data to improve decision making and business operations must make a concerted effort to hire and retain a diverse group of analysts with expertise in a variety of areas. These include strategy, data management, engineering, mathematics, statistics, programming, and cloud computing. In addition, analysts need to have proficiency in multiple computing platforms and analytic techniques, including predictive modeling, machine learning and data visualization, and must excel at clear, concise communication of insights.

Section 6: References

1. Utility-Scale Smart Meter Deployments: Building Block of the Electric Power Grid. Edison Foundation Institute for Electric Innovation, September 2014.
2. Program on Technology Innovation: Data Analytics and Customer Insights. EPRI, Palo Alto, CA: 2014. 3002003545.
3. "PG&E Ranks No. 1 in Smart Meter Deployments," News and Perspectives from Pacific Gas and Electric Company," September 19, 2014. Accessed online October 30, 2015: http://www.pgecurrents.com/2014/09/19/pge-ranks-no-1-in-smart-meter-deployments/.
4. "Company Profile," PG&E website. Accessed October 18, 2015: http://www.pge.com/en/about/company/profile/index.page?
5. "State Renewable Portfolio Standards and Goals," National Conference of State Legislatures website. Accessed October 30, 2015: http://www.ncsl.org/research/energy/renewable-portfolio-standards.aspx.
6. "Renewables Portfolio Standard," DSIRE website. Accessed November 11, 2015: http://programs.dsireusa.org/system/program/detail/840.
7. "What is a SmartMeter™?" PG&E website. Accessed November 1, 2015: http://www.pge.com/en/myhome/customerservice/smartmeter/index.page.
8. "About Teradata," Teradata website. Accessed November 2, 2015: http://www.teradata.com/about-us/?ICID=Awit&LangType=1033&LangSelect=true.
9. "The Basics of the SmartAC™ Program," PG&E website. Accessed November 2, 2015: http://www.pge.com/en/myhome/saveenergymoney/plans/smartac/basics/index.page.
10. How PG&E is Using Demand Management to Make Smarter Use of the Grid," PG&E Currents, October 30, 2014. Accessed online October 30, 2015: http://www.pgecurrents.com/2014/10/30/how-pge-is-using-demand-management-to-make-smarter-use-of-the-grid/.
11. Frequently Asked Questions about the Electric Program Investment Charge Program (EPIC), California Energy Commission website. Accessed online November 9, 2015: http://www.energy.ca.gov/research/epic/faq.html.

12. “Electric Program Investment Charge, PG&E website. Accessed October 19, 2015: http://www.pge.com/en/about/environment/pge/epic/index.page.

Export Control Restrictions

Access to and use of EPRI Intellectual Property is granted with the specific understanding and requirement that responsibility for ensuring full compliance with all applicable U.S. and foreign export laws and regulations is being undertaken by you and your company. This includes an obligation to ensure that any individual receiving access hereunder who is not a U.S. citizen or permanent U.S. resident is permitted access under applicable U.S. and foreign export laws and regulations. In the event you are uncertain whether you or your company may lawfully obtain access to this EPRI Intellectual Property, you acknowledge that it is your obligation to consult with your company's legal counsel to determine whether this access is lawful. Although EPRI may make available on a case-by-case basis an informal assessment of the applicable U.S. export classification for specific EPRI Intellectual Property, you and your company acknowledge that this assessment is solely for informational purposes and not for reliance purposes. You and your company acknowledge that it is still the obligation of you and your company to make your own assessment of the applicable U.S. export classification and ensure compliance accordingly. You and your company understand and acknowledge your obligations to make a prompt report to EPRI and the appropriate authorities regarding any access to or use of EPRI Intellectual Property hereunder that may be in violation of applicable U.S. or foreign export laws or regulations.

The Electric Power Research Institute, Inc. (EPRI, www.epri.com) conducts research and development relating to the generation, delivery and use of electricity for the benefit of the public. An independent, nonprofit organization, EPRI brings together its scientists and engineers as well as experts from academia and industry to help address challenges in electricity, including reliability, efficiency, affordability, health, safety and the environment. EPRI also provides technology, policy and economic analyses to drive long-range research and development planning, and supports research in emerging technologies. EPRI's members represent approximately 90 percent of the electricity generated and delivered in the United States, and international participation extends to more than 30 countries. EPRI's principal offices and laboratories are located in Palo Alto, Calif.; Charlotte, N.C.; Knoxville, Tenn.; and Lenox, Mass.

Together...Shaping the Future of Electricity

Programs:

End-Use Energy Efficiency and Demand Response

Understanding Electric Utility Customers

3002005720

Electric Power Research Institute
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com