PBS Turns to Google Analytics 360 and Google Cloud Platform to Deepen Audience Understanding

PBS

LunaMetrics®

About PBS

- PBS is a private, nonprofit corporation through which noncommercial, educational licensee members operate 353 PBS member stations.
- Headquarters: Arlington, VA
- pbs.org

About LunaMetrics

- LunaMetrics is a digital analytics and search marketing consultancy and a proud Google Analytics Partner since the beginning of the program.
- Headquarters: Pittsburgh, PA
- lunametrics.com

Goals

- Deepen understanding of audience segments
- Deploy tools to mine web data for insights
- Fine-tune site to meet segment needs

Approach

- Exported Google Analytics 360 data into BigQuery
- Used the Google Cloud Platform to interchange and analyze data sets
- Distilled data from 330 million sessions and 17.5 million episodes watched

Results

- Identified seven key audience segments based on web behaviors
- Developed in-depth personas per segment
- Informed content and feature development

Since its founding in 1969, PBS has reached millions of people with a wide array of television programming aired across 353 member stations. As an online content hub, PBS.org supports the broadcaster's television experience, striving to educate, entertain, and inform people from a wide variety of backgrounds and with a wide variety of needs.

Site visitors to PBS.org can find member stations, check local broadcast schedules, get news, discover new programs, watch clips or full episodes, or experience unique digital content based on a favorite show. As a Google Analytics 360 customer, PBS uses data to see which devices are used to visit the site, which pages are viewed, how long videos are played, and more.

But the sheer size of the PBS dataset is overwhelming. In a single year, PBS amasses data on over 330 million sessions, 800 million pageviews, and 17.5 million episode plays. Even armed with Analytics 360, a human analyst could spend weeks, or even years, trying to see patterns in a dataset that large. And while descriptive statistics such as average page views can help describe the big picture, they don't offer a deeper understanding of key audience segments. The team at PBS.org wanted more.

Dealing with really big data

"We've been saying for a long time that there is no 'average user' of PBS.org—rather that there are categories of users who do different things," said Amy Sample, PBS's senior director of strategic insights. "Differences in behavior are important in driving how we design and present content on the site."

Data mining—using computational algorithms to unearth hidden patterns in the data—could unlock shared behavior patterns to segment the PBS audience. In partnership with LunaMetrics, a Google Analytics 360 Services and Sales Partner, PBS embarked on a project to use a year's worth of data to do just that.

> *"This is the first time we've been able to use data mining across the totality of our web data to see what it can tell us about our audiences."*
>
> *—Amy Sample,*
> *PBS Senior Director,*
> *Strategic Insights*

Analytics 360 made it easy for PBS to move into data mining. Despite its massive size, the entire PBS dataset could be seamlessly exported from Analytics 360 into BigQuery, part of the Google Cloud Platform. BigQuery is designed to help customers query massive datasets by enabling super-fast, SQL queries—all while using the processing power of Google's infrastructure.

"We believe segments exist based on a variety of data sources, our contextual knowledge of the site, and of PBS viewers and members. But this is the first time we've been able to use data mining across the totality of our web data to see what it can tell us about our audiences," said Sample.

Taking analysis into the cloud

LunaMetrics selected 40 factors from Analytics 360 data to use in BigQuery. "Even with a tool as powerful as BigQuery, the first step was to pare down the dataset to make it manageable for analysis," said Jonathan Weber, data evangelist at LunaMetrics and advisor to PBS on this project. "BigQuery makes this easy and fast."

The team focused on page types (such as news, video, pages for individual programs, and content for parents), device types (desktop or mobile), traffic sources (links, search engines, and social media), and key interactions (logging in or clicking links to member station sites). From there, LunaMetrics and PBS used data science techniques to ultimately yield user groupings based on similarities in behavior.

"The datasets we were operating on were still very large," Weber said, "so we knew we needed to operate in the cloud." While the task was too big for a typical laptop, dedicated computing hardware didn't make sense. Google Cloud Storage was a handy storage space to inter-change data among services, and Google Cloud Datalab offered the right toolkit for analysis.

Cloud Datalab makes it possible to deploy Google Compute Engine resources in one click. This gave LunaMetrics a powerful environment for data analysis and visualization with an interactive format for the iterative development of code, charts, and documentation. "The Google Cloud Platform gave us the perfect tools. We could spin them up and down as needed, only paying for what we used," Weber explained.

> *"That was the coolest thing about it: A machine took our data without prior assumptions and reaffirmed and strengthened ideas that subject matter experts already suspected about our audiences based on our contextual knowledge."*
>
> **—Dan Haggerty,**
> **PBS Director, Digital Analytics**

Running with patterns revealed

Data mining algorithms can look at many criteria simultaneously to classify users by similarity. With data mining analysis, PBS was able to see patterns that would have been impossible to discern from averages.

Clusters in audience behavioral data revealed segments with widely varying needs, including groups who wanted to: 1) complete specific actions, such as activating PBS content on TV-connected devices; 2) focus primarily or even exclusively on one type of content, such as news articles; and 3) watch video in a preferred manner—via the site-wide video portal or individual show pages.

Dan Haggerty, director of digital analytics for PBS, said that the results made intuitive sense to the PBS team. "That was the coolest thing about it. A machine took our data without prior assumptions and reaffirmed and strengthened ideas that subject matter experts already suspected about our audiences based on our contextual knowledge."

More knowledge enables more nuance

With audience segment insights in place, the team at PBS can make more confident decisions about features and content to prioritize on PBS.org. PBS experts can query Analytics 360 data with the benefit of segmentation to target audiences for remarketing via AdWords or DoubleClick and for personalization using Google Optimize 360.

Rich audience insights are just the beginning for PBS. More data mining can help the broadcaster pursue content personalization by audience, tailor promoted content in sites and apps, and inform email updates and app push notifications across the entire PBS ecosystem.

Going forward, machine learning and human analysis will go hand-in-hand. Google Analytics 360, BigQuery, and the Google Cloud Platform enable the PBS team to glean insights from massive data sets and pursue more nuanced opportunities than ever before. "The most interesting lesson is that it's not an either/or proposition of 'machines' versus 'experts,' with one as a replacement for the other," said Haggerty. "Data science is really an interplay between the two."

About the Google Analytics 360 Suite

The Google Analytics 360 Suite offers powerful and integrated analytics solutions for today's biggest enterprises. Measure, see, and improve the impact of your marketing across every screen, channel and moment in today's customer journey. It's easy to use, and makes data accessible for everyone so the "aha" moments are simple to discover and share. Move from insight to impact faster with the Google Analytics 360 Suite, and as a result, make the most of every consumer connection. For more information, visit google.com/analytics/360-suite