

Viewpoint

David Menninger

SVP &
Research Director

Better Real Estate Insights with External Data

In an ever-changing real estate market, data and economic indicators are essential to formulating financial strategies. External data, therefore, is key. It is used by planning and operations teams to benchmark a firm's financial or operational performance. It enriches the data an organization collects and analyzes about customers, financial institutions, property managers, brokers and agents. It augments machine learning analyses to produce more accurate forecasting models. For both the residential and commercial sectors, demographic data is essential to understanding customers' buying behaviors.

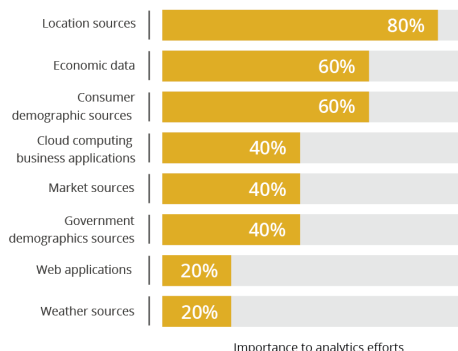
External data is an important part of real estate organizations' analytics efforts. The most important external data source identified in our research is location sources, followed by economic data. Organizations also identified consumer demographics, business applications, market sources, government demographics, web applications and weather sources as important to their analytics efforts. External data is not just part of machine learning analyses, though. Our research shows that external data sources are also a routine part of data preparation processes with all participating organizations incorporating at least two external data sources, with 60% of those organizations using more than 11 sources.

Real estate organizations face several key challenges when working with external data. First, they must decide which data to purchase, how to maximize the value of their spend and how to manage the ongoing budget associated with external data. Second, they must help individuals within the organization understand and find the data that has been acquired. Third, organizations need a strategy for storing and managing external data. And finally, as with any data asset, real estate organizations must maintain the external data to ensure it is up to date. Let's examine these challenges in turn.

Some external data sources are free, such as data published by governments and their agencies, and information on recent house sales. But data from commercial sources, such as demographic data, are often licensed by the volume of data requested, so budget concerns typically dictate that you know which data is needed. The process of selecting data to purchase can slow down the business since there can be lag time between when the data is selected, the purchase is approved and the data is

External Data Sources

Real estate organizations deem several important



Source: Ventana Research
Analytics and Data Benchmark Research
© 2023 Ventana Research



made available. In addition, departments can end up with runaway budgets as more and more external data is requested. Alternatively, organizations may buy all the data they anticipate will be needed instead of requiring departments to specify their data selections. This approach can, of course, be expensive as organizations may buy more data than they need.

If external data sources are not understood or if individuals cannot find the information they need for their analyses, data will go unused and the analyses will be less accurate. Real estate organizations can reduce this risk by deploying ways to catalog and search the information. With machine learning, organizations can go further. Introspection of the specific analysis being conducted, combined with information about the external data available can result in recommendations of which data to use and even recommendations on the type of analysis to perform. For instance, we assert that through 2025, four-fifths of property management organizations in the real estate industry will use digital technologies to streamline processes and align strategy to operational objectives.

One of the questions organizations face is how to manage this information. Many choose to load this information into their data lakes or data warehouses. This may seem like the obvious approach, but it also presents some challenges. The process of loading information may be slow, especially if incremental data purchases are made. Loading the incremental purchases should be reviewed and approved, even if an automated routine has been created to handle the incremental data. External data can often contain sensitive information subject to privacy or other policy restrictions and the governance of that data must be considered as incremental purchases are made.

Processes must also be in place to maintain the data. Market data is constantly in motion, for example, as customer demographics change or as broker contract histories grow. Various economic indicators fluctuate with expanding and contracting cycles. Maintaining up-to-date information is critical to ensuring accurate analyses and supporting optimal operational decisions in the real estate industry.

One way to address these issues is with an external data platform or “data as a service.” The external data platform provider manages the information, keeps it current and makes it available as necessary. In some cases, the platform can assist with analyses by automatically discovering features of the data and their potential impacts on machine learning models. The platform can also help match and validate multiple external data sources with each other and with internal data sources. Where data needs to be loaded into internal systems, the data vendor can provide various ways to export data or analyses such as machine learning models based on the data.

Real estate organizations should examine how better managing external data could enhance their financial, operational and analytical processes. As the number of data sources grows, they should address the issues associated with purchasing, storing, accessing and maintaining this data. Each real estate organization will need to address these issues with approaches appropriate for its situation, but each should ensure that its strategy encourages a well-managed use of external data.



Dave Menninger – SVP and Research Director, Ventana Research

David Menninger is responsible for the overall direction of research on data and analytics technologies at Ventana Research. He covers major areas including artificial learning and machine learning, big data, business intelligence, collaboration, data science and information management along with the additional specific research categories including blockchain, data governance, data lakes, data preparation, embedded analytics, natural language processing (NLP) and IoT.