



Data Science Case Study

Community Healthcare Utilization Risk Score

Mosaic helped a leading hospital system change the way they deliver care by understanding their communities through machine learning.



Industry

Healthcare



Use Case

Population Health



Techniques

Unsupervised Learning,
Hierarchical Regression,
AI Adoption



Outcome

Uncovered hidden trends on the hospital's population and provided actionable information to decision-makers

WHY POPULATION HEALTH ANALYTICS MATTERS

As healthcare organizations develop more sophisticated analytics capabilities, they are beginning to move from descriptive analytics and reporting towards the world of causal inference, machine learning, predictive analytics, and optimization. This leap to advanced analytics is a major change of perspective for any organization and requires the right skill-sets, data, organizational openness, and culture to take full advantage of the tools available to businesses today.

The healthcare industry, hospitals in particular, offer a number of improvements ripe for machine learning and predictive analytics. One of these areas is population health management. Hospitals can change the way they deliver care by understanding their communities through the application of machine learning and statistics. Once healthcare systems understand the drivers of utilization rates, they can begin to cluster community segments and ultimately implement preventative interventions in the communities they serve.

A HOSPITAL CONTINUES TO INNOVATE

One of the Nation's top ranked and most innovative Children's Healthcare systems reached out to Mosaic for healthcare machine learning support. The hospital is a pediatric healthcare network devoted to the health and well-being of children. The hospital is known both for its nationally and internationally recognized medical, research, and education programs and its full range of everyday care for kids throughout their own and surrounding states. With more than 1,000 healthcare professionals representing the full spectrum of pediatric specialties, the hospital's network includes its main campus, more than a dozen locations and hundreds of outreach clinics.

The hospital has a history of innovation that is not just medical but even extends to data & analytics. For example, the hospital developed, validated, and deployed its own custom patient risk score, known as the "Patient Complexity Score" (PCS). This score leverages electronic medical records and other patient data such as patient responses to social determinants of health screener questions. Clinicians use the PCS to quickly understand how much care a child might need, thereby improving care decisions.

Moreover, the hospital is a non-profit that is committed to promoting the health of children in its local community. This commitment is evidenced by the health advocacy institute the hospital runs, which collaborates with the public and community institutions to deploy evidence-based approaches that improve the health and safety of local children. The institute understands that to provide the best healthcare for children, healthcare institutions and professionals must take into account children's social and environmental context and even work to improve it — in a sense, the community is the patient.



COMMUNITY HEALTHCARE UTILIZATION DISPARITIES

The hospital noticed that patients living in three zip codes close to the main campus seemed to utilize healthcare resources much more than an average patient. Furthermore, they tended to utilize complex and high-cost hospital resources. Healthcare leadership suspected that they could more effectively serve these communities and patients by providing clinicians with additional relevant contextual information about their home communities and through interventions like greater investments in community partners and programs. Furthermore, our customer, like many hospitals, is affected by the transition from fee-for-service to outcome-based policies among healthcare payers, and is thus motivated to learn how to best promote health among its patients, no matter their zip code.

Finding empirical evidence to guide the health system's efforts to better serve its local community would require the integration of internal data describing patients and their health and healthcare with external data describing local communities, along with sophisticated statistical modeling to identify relevant community-level drivers of higher healthcare utilization. The healthcare provider enlisted Mosaic, a leader in healthcare statistical and machine learning consulting, to work with them on this analysis effort, as well as to help them grow their in-house data science and analytics capabilities.

DEVELOPMENT OF A RISK SCORE TO IMPROVE DECISION MAKING

Our hospital client and a dedicated population health analytics team decided to build upon the hospital's success with their in-house Patient Complexity Score (PCS) by co-developing an analogous Community Complexity Score (CCS). The CCS is an interpretable score for a community specified by a geographic area that quantifies the risk that an individual living in that community will utilize relatively high levels of healthcare, after adjusting for individual characteristics such as PCS and age. The CCS is intended to be used both in clinical settings for decisions related to individual patient care and in community outreach settings for decisions related to community organization partnership or wider community-based interventions.

A diverse team including Mosaic data scientists and hospital subject matter experts and data & analytics professionals was assembled to develop the CCS. The process kicked off with reviewing relevant literature, identifying and exploring promising external data sets, and combining these external data sets with internal data sources such as historical electronic medical records. Exploratory data analysis informed statistical and predictive modeling and feature selection, which in turn led to the specification of the algorithm for computing a proof-of-concept CCS.



Mosaic consultants identified an advanced statistical modeling technique known as [hierarchical or multilevel regression](#)¹ as a great fit for this challenge. This technique allows for simultaneous analysis of the impact on a target variable of both characteristics of individuals and characteristics of communities at a higher “level” than the individuals from those communities, and assumes that these communities are related (i.e. their influence on the outcome is governed by an overarching distribution). The consultants then developed a variant of a feature selection approach to select community-level features that were improving the fit of the model to the data set of individual characteristics, their healthcare utilization, and features of their home communities.

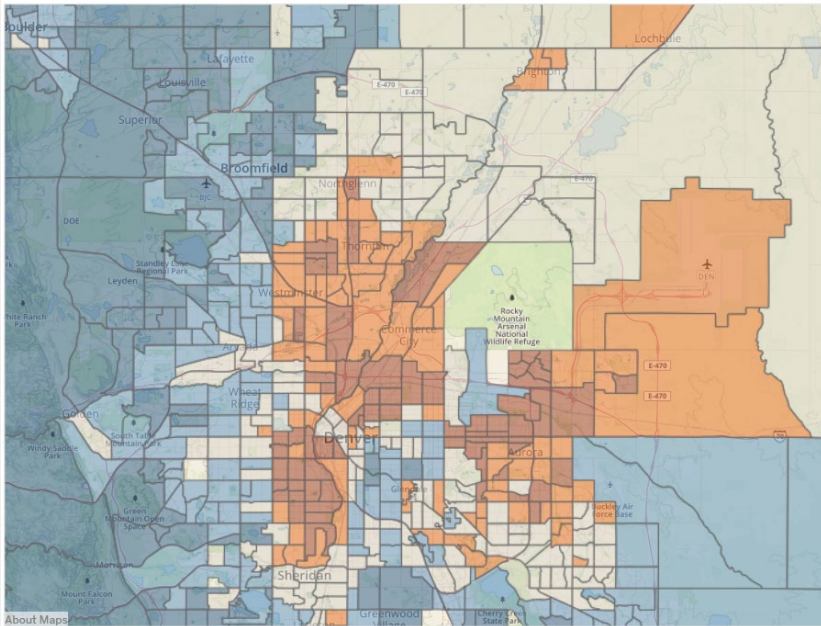
Community-level features would only be selected by this process if they were able to explain variation in the target healthcare utilization variable after accounting for individual or encounter characteristics. This gave the hospital confidence that it was finding community-level drivers of higher healthcare utilization to include in the CCS. These selected features and their weights in the resulting models were used as the basis for the CCS computation algorithm. Ongoing work with the hospital seeks to synthesize the results of the hierarchical/multilevel regression feature selection approach with other predictive modeling and feature selection approaches, as well as hospital subject-matter expertise. Once this effort leads to a new set of features, unsupervised learning techniques like Principal

Components Analysis or Factor Analysis will be used to develop a new set of weights for an updated CCS.

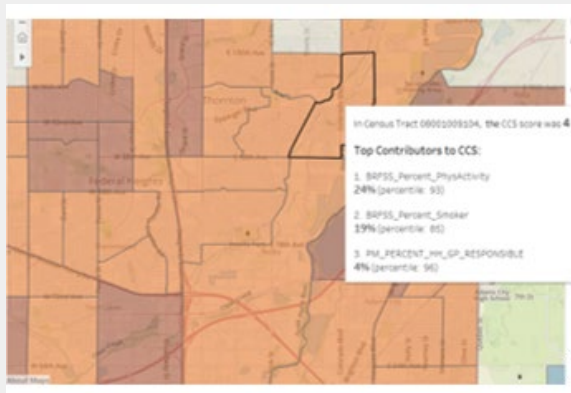
Within just a few months, the team developed a proof-of-concept CCS, created an early prototype of a CCS delivery application (screenshots below), and collected feedback from potential CCS users to guide further development. Both clinical and community outreach decision makers were enthusiastic about the potential for the CCS to improve their decisions. They appreciated its interpretability and suggested ways that it might most effectively be delivered to them along with other relevant information. The team quickly progressed through critical stages of an iterative development process, creating greater clarity regarding the most promising next steps.



Community Complexity Score by Census Tract

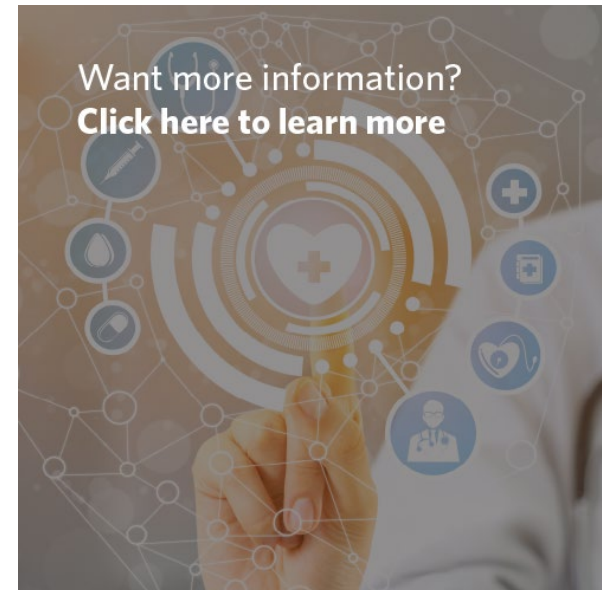


Population Health Analytics Visualization



BENEFITS

Mosaic has worked and continues to work closely in partnership with the provider's data & analytics team members on CCS development and other projects. The hospital is finding value while growing from descriptive into diagnostic analytics that uncover causes and motivate evidence-based interventions. Hospital administrators have provided their feedback on the prototype community score, a prototype Tableau dashboard that delivers the CCS to community health navigators has been deployed, and the next round of improvements and analytics is currently being developed as additional feedback is collected from dashboard users.



Endnotes

1. https://books.google.com.pr/books?hl=en&lr=&id=c9xLKzZWoz4C&oi=fnd&pg=PR17&dq=hierarchical+or+multilevel+regression&ots=bbS3L2Nsrd&sig=MCepZv_NOWxSxRiNjtbFABI-UIA&redir_esc=y#v=onepage&q=hierarchical%20or%20multilevel%20regression&f=false

