Using Teradata Business Analytics Solutions to Fight Cancer

06.17 EB9766 LIFE SCIENCES

TERADATA.

Advanced analytics have enabled businesses to know customers on a personal level and customize offers to meet their individual needs. Analytics are now enabling that same type of personalization in medicine.

Life sciences are using data and analytic insights to understand a person's genome makeup to identify genomic patterns that could indicate cancer.

The Power of Genome Data

The 1000 Genomes Project sequenced the genomes of 1000 people from around the world to create the most detailed and medically-useful catalogue of human genetic variation. It helped researchers discover more than 100 regions of the genome that contain genetic variants associated with common diseases such as diabetes, coronary artery disease, prostate and breast cancer, rheumatoid arthritis, inflammatory bowel disease, and age-related macular degeneration.

Data from the 1000 Genomes Project is now being used to show genetic similarities and variations across human demographics and geographies. The life sciences industry plans to use the data to unlock medical secrets that will lead to more effective treatments, including the treatment of cancer.

This brings doctors one step closer to the Holy Grail of science and medicine—the ability to deliver specialized treatment for each individual patient.

Knowing a person's particular genome composition allows more effective treatments and can lead to earlier cancer detection. That's because each individual's genome history and chromosomes influence their health and response to medicines. When doctors have this information, they can prescribe more effective treatments, including treating diseases before any symptoms are evident.

Visualizing Complex Genomic Data

Teradata and VizExplorer™ brought together a diverse group of data scientists, genomic doctors, cancer researchers, technologists and data visualization experts to work on a genome project. Their mission entailed understanding more than 1000 genetic patterns, which could pave the way for individualized treatments based on a person's genetic makeup.

Data from more than 75 trillion records provided details about 25,000 tumors, which equated to more than one trillion data points. This massive amount of data was integrated with open data from the 1000 Genome Project in a powerful and scalable data warehouse with a Teradata® Database. The team was challenged to uncover insights in the incredibly high-dimensional data set and present the findings in a way that doctors, scientists, and researchers could readily understand.

Developing an analytic model and algorithm for the complex genomic data was an extremely complicated task. After experimenting with various analytic techniques, the team realized that Quartal Super Graphics from VizExplorer proved to be the most effective. The solution applied an in-database, recursive algorithm that processed the entire 1000-genome-population into a common, hierarchical Quartile grid. A database query, capable of large-scale processing, then built a subset of data for each of the corresponding genome communities within the total population. The data subset was then used to create heat maps that referenced specific genomes and visualized the information.

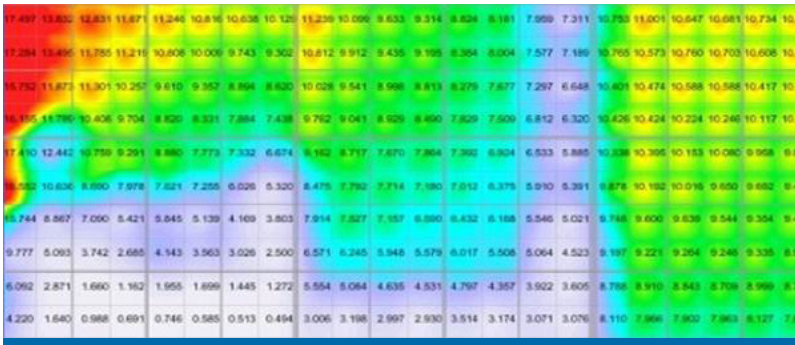
The resulting graphics allowed the sequence of data to be seen across communities for the entire 1000 Genome Project. This showed, for example, how genomes that originated in China were different from those in Africa and from the United States. Deeper analysis of the data found more granular differences and similarities. This valuable information can be used to determine how certain genomes are susceptible to cancer.

Decoding Genomes Enables Insights into Cancer

Using data and analytics to decode the human genome accelerates the development of the treatments and medication to treat diseases like cancer. The analysis helps identify cause and effect relationships, correlations, and links, which are vital keys to unlock disease management and cures.

Data from the genomes provides insights to answer several pressing questions:

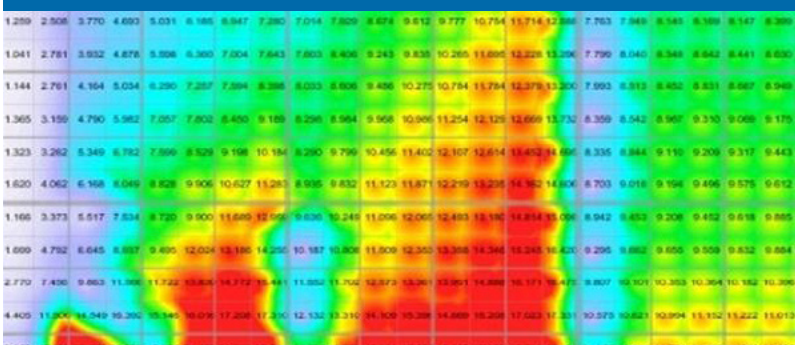
- Which cancers are impacting humans the most?
- Where did the single common genome cancer originate?
- Which genome patterns indicate a susceptibility to cancer?
- How can genome patterns be used to develop personalized treatments?
- How do genomes compare in people in different geographic areas?



Visual Insights into Genetic Patterns

Teradata and VizExplorer worked together to visualize genome data and identify patterns of cancer. The result was "Genome World Window," which are Art of Analytics images that show a community or geography within the 1000 Genome Project. Each artwork, or frame, is a heat map showing genome characteristics in people that may cause cancer. Comparing the artwork for each geography highlights the genomic variations among communities across the globe.

The Art of Analytics images show the significant value of large-scale genome analytics and high-intensity graphic technologies. The art allows doctors, other medical professionals, and even patients to easily see genetic patterns of cancer. A better understanding of these patterns will lead to personalized medical treatments based on each person's genetic composition.



The genome analysis was able to map the development and movement of historical migration patterns to show how demographics correspond to social aspects in humans. Seeing how migration happened over centuries provided insights into how pure genomes, when merged with others, affect cancer. Data showing when humans moved across the globe can also show how instances of cancer changed as genomes were mixed.

Contextualizing Data to Solve Problems

Teradata Business Analytics Solutions can help any business and industry use the same approach as the genome project to bring together highly skilled experts, including data scientists, and use their knowledge to solve a complex challenge. People, products, processes, services, and technology can align toward a common goal to improve businesses or, like with the genome project, advance life sciences to improve peoples' quality of life.

Teradata can bring together the right people and technologies to address any issue. With an analytic approach that gives context to data, organizations can answer difficult, otherwise elusive questions through new insights. As companies are being flooded with increasing volumes of data from a growing number of sources, it's more important than ever for businesses to have solutions capable of analyzing large and complex data sets. Whether it's gaining a deeper understanding of cancer or solving another pain point, Teradata can help companies achieve the high-impact outcomes needed to advance their business.

Connect with Teradata

Watch our video series, The Art of Analytics, featuring Genome World Window for more details on ways Teradata data scientists and consultants work with massively large data sets to find medical insights that will result in a personalize approach to treatments. To find out more about how business analytics solutions from Teradata can deliver high-impact business outcomes, visit Teradata.com.

10000 Innovation Drive, Dayton, OH 45342 **Teradata.com**

06.17 EB9766