2019 NPCR COLORADO SUCCESS STORY

Colorado Cancer Registry: John Arend and Kirk Bol

Using Tableau to calculate and display customizable cancer data

NATIONAL PROGRAM OF CANCER REGISTRIES
SUCCESS STORY

SUMMARY: Making cancer registry data available to the public is a core function of a surveillance registry. Colorado Central Cancer Registry (CCCR) staff assisted the Colorado Department of Public Health and Environment (CDPHE) in the creation of a new public-facing Tableau dashboard, a re-imagining and re-engineering of the Colorado Health Information Dataset (CoHID), the premier of accessing Colorado's public health surveillance data. This tool allows researchers and the public to create custom data visualizations from a selection of data sources, including cancer registry incidence and mortality data; meanwhile saving staff time and providing data to the public in a visually-appealing format.

CHALLENGE: In the early 2000s, CDPHE developed and launched the Colorado Health Information Dataset (COHID), which at the time was among the first web-based, user-searchable platforms for access to public health surveillance data. COHID originally comprised modules for accessing cancer registry data, vital statistics (birth and death), birth defects and population data. It was further enhanced to include data from Colorado's Behavioral Risk Factor Surveillance System (BRFSS), Pregnancy Risk Assessment Monitoring System (PRAMS) and other data sources. Among the goals of COHID was to provide data users a single point-of-entry for accessing key data to inform public health efforts at the state and local level. While cutting-edge in its time, the underlying architecture of this first-generation COHID had reached end-of-life, in terms of both software and hardware. Additionally, while robust in its ability to stratify data, its simple, tabular output lacked visual appeal. Many states grapple with how to provide user-friendly, customizable and digestible data visualizations for multiple types of data to the public, which typically requires contracting out that service or trying to develop internally, both of which cost valuable resources. Further, off-the shelf solutions, while often more affordable, rarely deliver the ease-of-use promised nor the opportunity for customization required by the nuances associated with various public health surveillance data.

SOLUTION: Tableau provides a robust data visualization tool; however, prior use of its desktop software and reliance on Tableau's hosting of underlying data necessitated pre-calculating the desired data, creating numerous tables representing myriad combinations of cancer site/race/sex/county/year aggregations that require time and effort to create and recreate, minimizing the time-saving benefits of a public-facing reporting tool. With Colorado's investment in Tableau as an enterprise software solution, and CDPHE's purchase of a Tableau Server license and associated hardware, CDPHE staff were able to create a visualization solution that afforded local hosting of individual level data, and

the processing power to support concurrent use of multiple visualization across multiple datasets. More specifically, the new environment—retaining the COHID moniker--is a robust visualization that calculates crude and age-adjusted rates, with 95% confidence intervals and counts on-the-fly from record-level, individual data for a diverse array of stratifications. The use of Tableau Server keeps the record-level data stored within our secure network environment at all times while still allowing public access to the resulting statistics. CCCR staff invested extensive time upfront in refining and testing the view, verifying that the underlying algorithms were accurate, and ensuring that rules to maintain confidentiality were appropriately instituted. However, following this up-front investment of time and other resources, ongoing development of new features and ensuring routine updates to the data contained within COHID is made relatively easy, which should support a long life and ability to respond to emerging public health data needs.
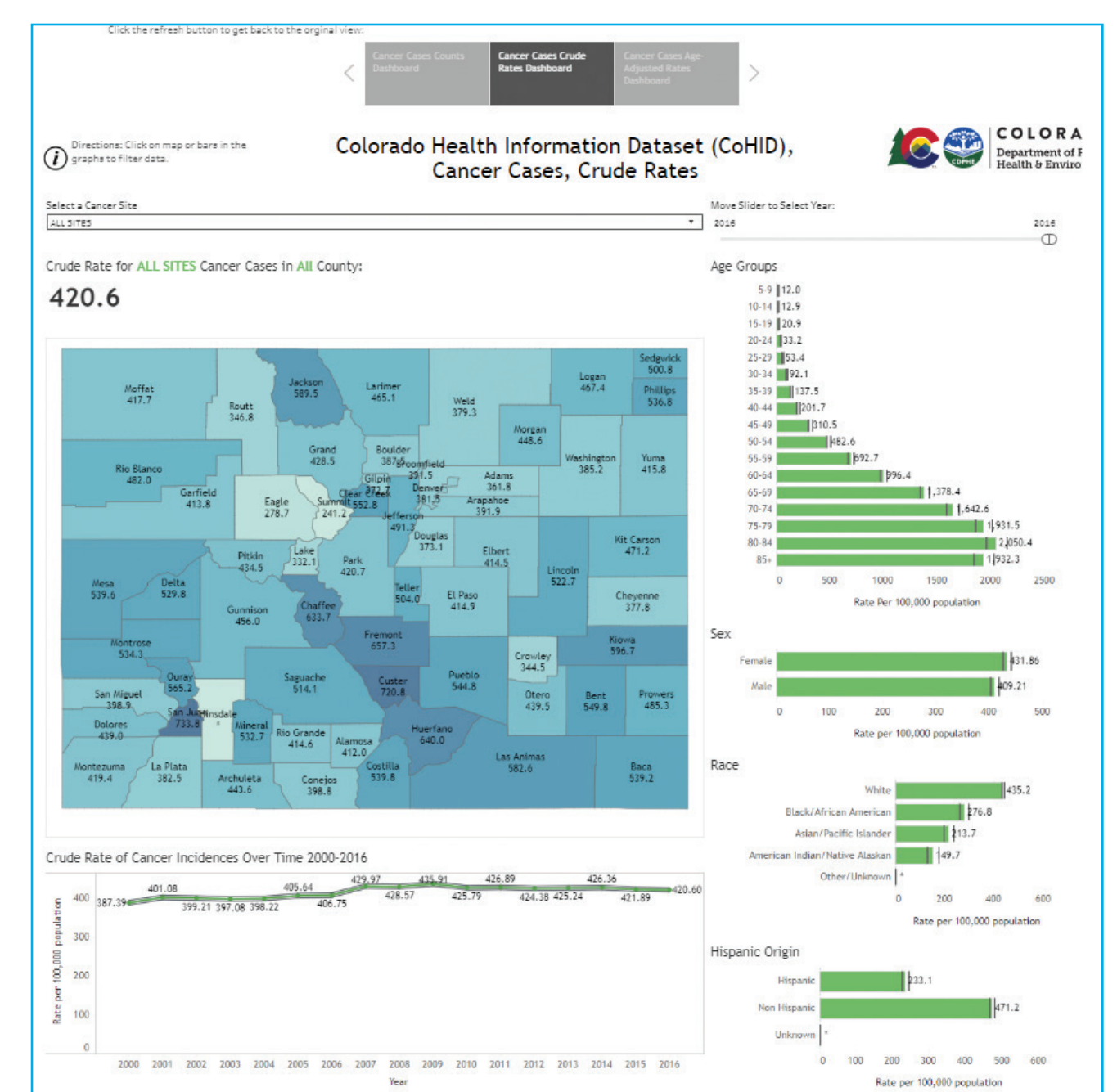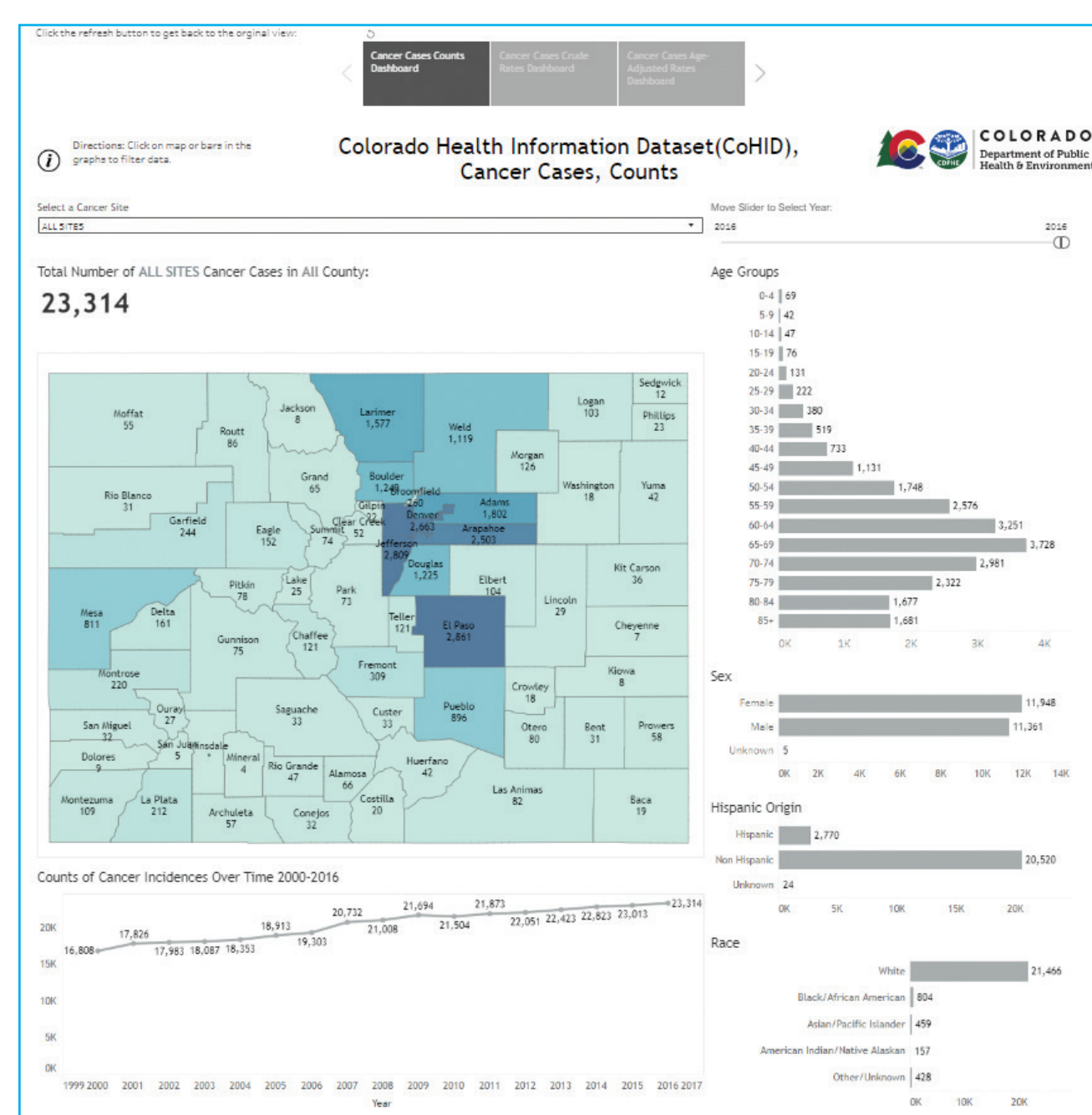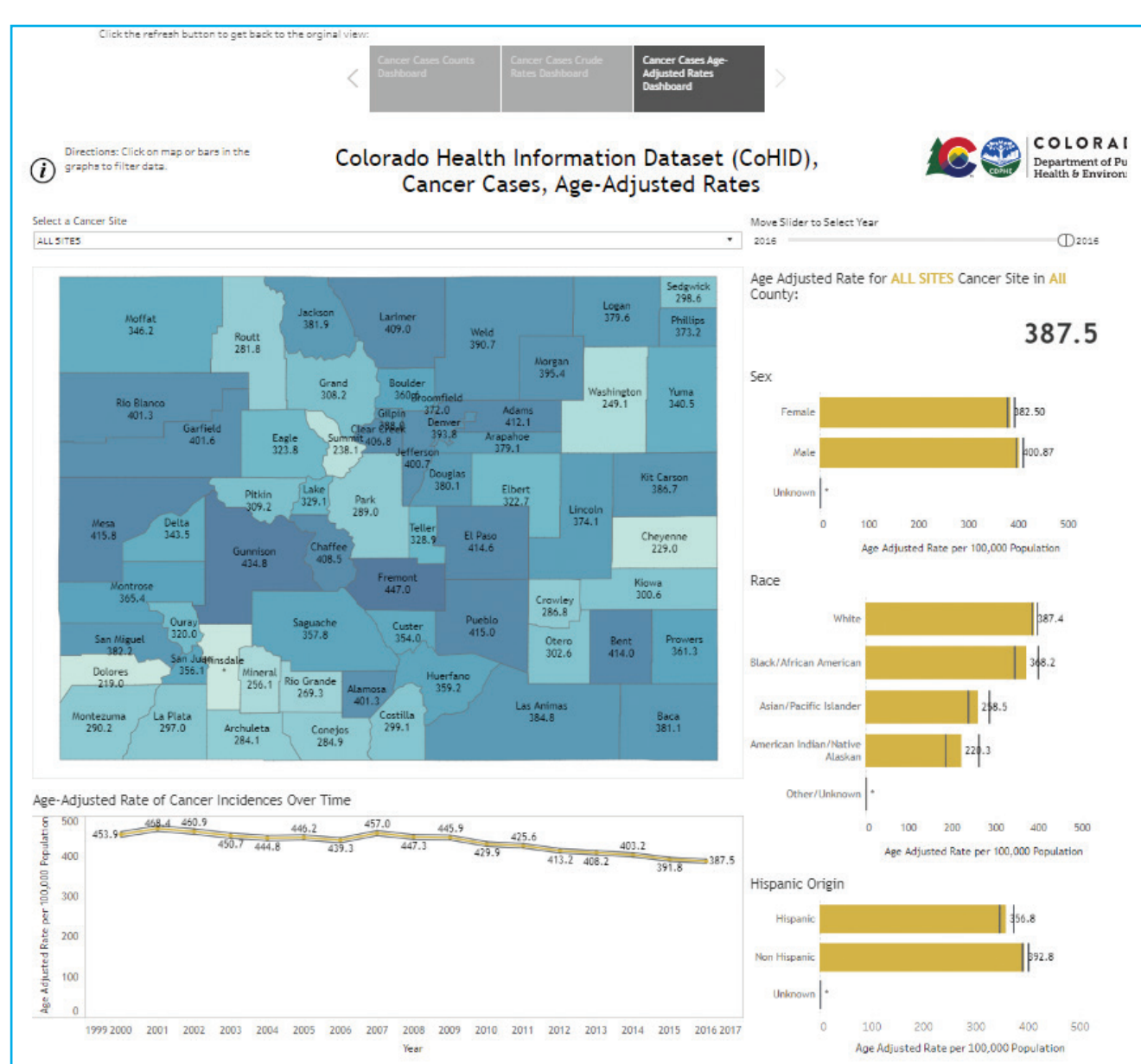
RESULTS: CDPHE's new CoHID site went live on August 23rd, 2019, hosting data on cancer incidence, birth, death, birth defects, suicide, drug overdose, population estimates, environmental epidemiology data, and more. By collaborating with other health department programs, staff time for development and testing was dispersed across the department and costs were minimized. All statistics are calculated in real-time to reflect the choices made by the user. The view includes a map of all Colorado counties and data are readily presented through hover tooltips. The CoHID cancer dashboard allows users to create their own custom stratifications by year, sex, race, cancer site and county. Users can combine one or more counties to create their own custom regions, can select one or multiple years, and can combine 5-year age groups. The new CoHID saves registry staff time as data requests can be easily referred to the online tool and updating the data to add additional years is a simple data extract. Early feedback from users has been very positive.

SUSTAINING SUCCESS: As CDPHE already budgets for Tableau Server, sustaining the new CoHID site requires minimal additional resources. Data updates require little more than extracting a new file as all statistics are calculated on the fly within Tableau. Multiple programs rely on the tool to make their data available to the public which increases the momentum to maintain and improve upon the existing tool. The CCCR continues to evaluate additional statistics and displays that could be produced and presented through CoHID.

CONTACT INFORMATION:

Tel.: 303-692-2540
https://www.colorado.gov/pacific/cdphe/cancerregistry