

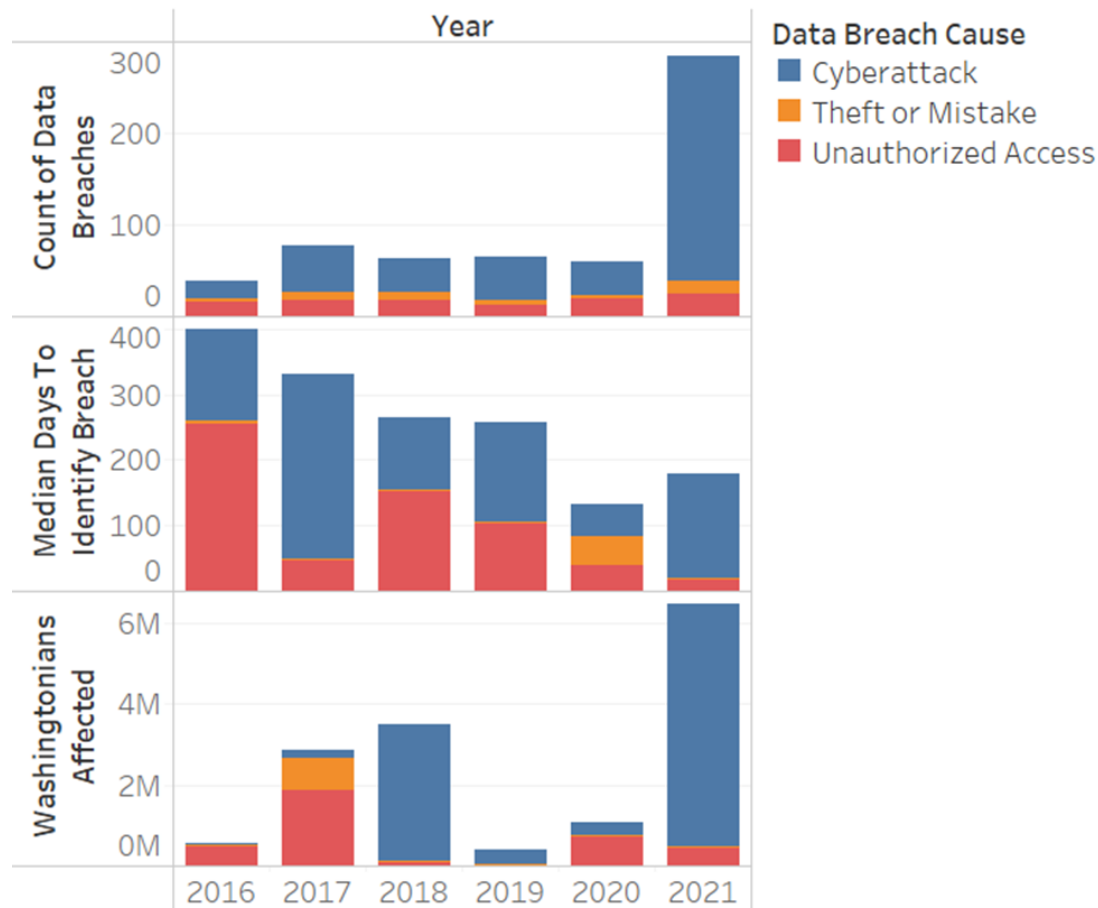
Autumn Erickson

November 11, 2021

Assignment #3: Data Visualizations

These graphs were created using data collected by the Washington State Attorney General's Office and include data on over 600 data breaches that occurred from 2016-2021 that affected at least 500 Washingtonians each. These visualizations are intended for cybersecurity professionals, specifically those who do business in Washington.

Data Breach Trends In Washington State 2016-2021

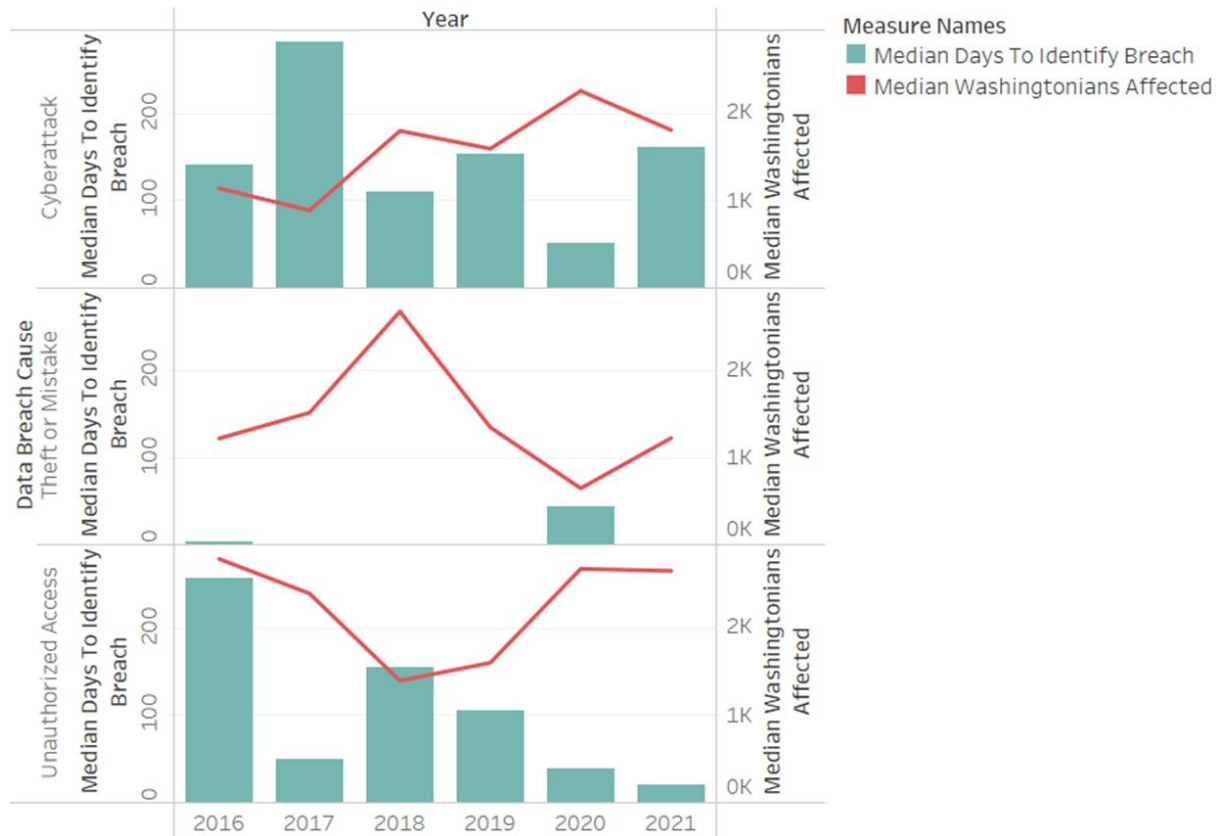


Count of data breaches, median days to identify each breach and sum of Washingtonians affected for each year. Color shows cause of data breach.

For this visualization I created three stacked bar graphs to measure total data breaches, median days to identify them, and total Washingtonians affected by year and cause of data breach. I selected these variables from the dataset because they are relevant to the entire audience, since there were also sector specific variables, and some of the other variables had missing data. This visualization is important for cybersecurity professionals because it not only shows the dramatic increase in cyberattacks seen so far this year, but also how they compare to

the effects of and responses to data breaches by other causes. I chose to use a stacked bar graph so I could show the broader trends in data breach impacts and responses, while also allowing the reader to see which data breach cause these trends can be attributed to. For example, in the “Median Days to Identify Breach” graph, we see that breaches are being identified quicker, and looking at the separations we then see that this is mainly attributed to the “unauthorized access” breaches being identified quicker, while the other two show no clear trends. This simple but informative visualization gives cybersecurity professionals a broad understanding of data breach trends in Washington, allowing them to make an informed response within their corporation. The next visualization, however, misleads them into a less informed and potentially incorrect response.

Data Breach Trends in Washington State 2016-2021



The trends of median days to identify breach and median Washingtonians affected for year broken down by data breach cause. Bars show median days to identify a breach, and lines show median Washingtonians affected per breach.

For this visualization I made combination charts separated by the data breach cause and plotted the median days to identify the breach, and median Washingtonians affected per breach for each cause by year. This visualization would only be useful for cybersecurity professionals if they wanted to see the level of threat posed by the average breach depending on its cause over those years. While this visualization provides accurate data, it is not as useful for the intended audience compared to the previous one, and can be misleading. Taking the median number of people affected per attack is not very affective because there are no clear trends, and the number

tends to stay within the range of 1,000 to 3,000 people. Since the visualization is meant for cybersecurity professionals, it is more important to show the total number of people affected per year, so they know the full effect of data breaches. If the reader had not seen the first visualization, they would not know about the increased prevalence and impact of cyberattacks in 2021, especially since this visualization shows that the median people affected per cyberattack actually decreased since last year. These graphs may be misleading because the reader may see “unauthorized access” breaches as the greatest threat since more people are affected per breach. While unauthorized access breaches are dangerous, the reader does not know the scale of the issue and that only a fraction of data breaches and people affected by them are due to unauthorized access. There may be further confusion since the two variables being plotted together may imply there is a correlation between the two, but no correlation is seen. In addition, using a line graph is misleading since the data is by year, but the reader might follow the line to make estimates for specific months/quarters, which would be inaccurate. Lastly, the color choice may make reading these graphs difficult for colorblind people, especially where the lines and bars overlap.

Data Source:

Imamura, Kaya. Washington State Attorney General’s Office, *Data Breach Notifications Affecting Washington Residents*. Olympia, Washington. Distributed by Data.gov.

Updated November 7, 2021. <https://catalog.data.gov/dataset/data-breach-notifications-affecting-washington-residents>.