COVID Learning Data Overview

<https://www.kaggle.com/c/learnplatform-covid19-impact-on-digital-learning/data>

The data being examined comes from a variety of public and automatically collected data sources. The engagement data is sourced from LearnPlatform’s Student Chrome Extension, which collects page load events of over 10,000 education technology products in LearnPlatform’s product library, including websites, apps, web apps, software programs, extensions, ebooks, hardwares, and services used in education institutions. District data comes from a collection of information from the National Center for Education Statistics, The Federal Communications Commission, and Edunomics Lab. [1]

The engagement data is aggregated as school district level, with each file containing information on the time when the product was accessed, the unique identified of the product, the percentage of students in the district who had at least one page-load event of a given product and on a given day, and the total page-load events per one thousand students of a given product and on a given day. The district information data underwent an anonymization process to remove identifiable information about the schools districts and contains information regarding district id, state, locale, percent of students identified as Black or Hispanic, percent of students on free or reduced lunch, information regarding internet connection speed, and the per-pupil total expenditure. Finally, our product information data includes information about the top 372 products with most users in 2020 and contains information regarding the unique identifier of the product, the URL to the product, the product name, the product provider, the sector of education in which the product is used, and the product’s primary essential function. In summary of the data being examined, there is a total of 236 files and 945 columns. [1]

With this data, we hope to explore overall trends in the popularity of online educational tools, as well as how this has changed over time throughout the COVID-19 pandemic. Further, we hope to draw out patterns that may exist between the use of paid products and the percentage of students on free or reduced lunch in the district, thus attempting to identify a correlation between the quality of educational products being used and the general wealth of the district.