This directory contains data and code that replicates tables and figures for the following paper:

Title: The COVID-19 Pandemic Disrupted Both School Bullying and Cyberbullying

Authors: Andrew Bacher-Hicks, Joshua Goodman, Jennifer Greif Green, Melissa K. Holt

There are no specific computational requirements other than those necessary to run STATA. The authors were able to run all scripts in the following computing environment: Intel® Core™ i7-10850H CPU @ 2.70GHz (6 cores, 12 threads), 64 GB RAM; STATA/MP 17. The runtime was less than fifteen minutes.

Directory structure:

The code assumes the following folder structure:

> \code : contains all do-files necessary to generate the analysis file, create tables, and figures

> \data : contains the raw data and analysis file

> \paper : the provided Stata code will save figures and tables it generates

Software requirements:

STATA/MP 17 was used to run all do-files. The following non-standard Stata packages are called in individual do-files:

- reghdfe: for two-way fixed effects and clustering

- ftools: this package is required to run reghdfe

- coefplot: for plotting regression coefficients

- statastates: generates state abbreviations based on names

- _gwtmean: used to generate weighted means

Instructions:

1. Before running any of the do-files, set global initial path in 01_clean_data.do (line 5) and the two global paths in 02_run_analysis.do (lines 7 and 8).

2. Run 00_config_stata.do. This will install the five non-standard Stata packages referenced above.

3. Run 01_clean_data.do. This processes the raw data and generates the analysis file (analysis_file.dta) used to create all figures and tables.

4. Run 02_run_analysis.do. This file loads analysis_file.dta and produces all figures and tables for the paper.

Datasets:

1) **Youth Risk Behavior Survey (YRBS)**

   The YRBS data are used to provide national and state-specific estimates of school and cyberbullying. All data are publicly available online here: https://www.cdc.gov/healthyyouth/data/yrbs/data.htm.

2) **U.S. Census Stata Population**

   The state population data from the U.S. Census was used to generate weights for our national regression analyses. The data come from the following dataset "Single Year of Age and Sex Population Estimates: April 1, 2010 to July 1, 2019 - CIVILIAN (SC-EST2019-AGESEX-CIV)". The dataset is publicly available online here: https://www.census.gov/data/datasets/time-series/demo/popest/2010s-state-detail.html.

3) **Google Trends**

   The Google Trends data provide national and state-specific estimates of search intensity for specific topics and keywords related to school bullying and cyberbullying. These measures are publicly available and can be downloaded here: https://trends.google.com/trends/.

4) **Burbio**

   The Burbio data provide estimates of school learning modality. Weekly measures by school district are available for purchase here: https://cai.burbio.com/school-opening-tracker/. In this paper, we only use aggregate state-wide measures from September 2020 through February 2021, which we obtained permission to post publicly to allow for replication of this paper.