**Overview**

The NYT covid data is useful for tracking cases, but there is much more information on each institution which is why the IPEDS dataset was added. NYT already had this column, so that made it possible to do a join on the NYT dataset. During the merge, the IPEDS Illinois dataset lost 204 rows, since NYT is only tracking 52 colleges in Illinois. This sheds some light on how many colleges aren’t even reporting covid statistics.

**Date Windows**

In order to use static data, the NYT dataset was “frozen” on February 13 2021. NYT pushes to their github every 2 weeks, but in order to have consistent data analysis, we must freeze the data at some point.

The IPEDS data set reports on the 2019-2020 school year since the current school year is still being generated. This means that when we report on school demographics, we’re looking at snapshot of when the pandemic first struck, rather than the “aftermath” era we’re currently in. This means trends like the decline of international students at studying form home won’t be fully reflected.